

**Univerzita Karlova v Praze**

**Filozofická fakulta**

*Ústav Českého národního korpusu*

*Matematická lingvistika*

## **Disertační práce**

*Srovnávací aspekty lotyšského a českého lexikonu:*

*Materiály k sestavení lotyšsko-českého slovníku*

*Comparative aspects of Latvian and Czech lexicons:*

*Materials for assembling a Latvian-Czech dictionary*

vedoucí práce: Prof. PhDr. František Čermák, DrSc.

konzultant: Mgr. Pavel Štoll, Ph.D.

2016

*Michal Škrabal*



**Prohlášení:**

Prohlašuji, že jsem disertační práci napsal samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

*V Praze, dne 29. 3. 2016*

.....

*Jméno a příjmení*



### **Poděkování:**

Autorovou milou povinností je poděkovat na tomto místě mnoha lidem, bez jejichž laskavé podpory a rad by tato disertační práce byla jen stěží vznikla. V první řadě bych rád poděkoval prof. Františku Čermákovi, jenž svou roli školitele splnil bezvadně a zvláště v kritické finální fázi mi pomohl vybědnout z marastu podružností a apriorismů. Velikou vděčnost cítím vůči svému konzultantovi Pavlu Štollovi, jemuž bylo snášet můj lexikografický teror ze všech nejvíce: naše společné „slovníkářské dílničky“ se mu z paměti jistě jen tak neztratí, ztratí-li se mu vůbec kdy. Dále chci poděkovat jmenovitě doc. Marii Vachkové za to, že nade mnou (či rovnou o mne) nezlomila hůl, lotyšským kolegům prof. Jurisi Baldunčiksovi, prof. Ojārsu Bušsovi a Santē Jēraneové za vstřícnost při konzultacích a přátelům Gregu S. Evansovi a Jānisi Krastiņšovi za zprostředkování trestě mé práce případným nečeským zájemcům (Jānisovi pak nadto za bezpočet konzultací).

V neposlední řadě bych rád vyjádřil vděk své rodině za jejich vytrvalou podporu po čas mého – notně protahovaného – studia a všem kolegům z Ústavu Českého národního korpusu, ať už jsem po nich v posledních týdnech a měsících něco chtěl anebo ne.

## **Abstrakt**

**Název práce:** Lotyšský lexikon z hlediska dvoujazyčné lexikografie: Teoretický podklad pro tvorbu lotyšsko-českého slovníku

**Autor:** Mgr. Michal Škrabal

**Katedra:** Ústav Českého národního korpusu

**Školitel:** prof. PhDr. František Čermák, DrSc.

Primárním cílem práce je rozčlenit lotyšský lexikon, respektive jeho relevantní segment, do jednotlivých skupin, vymežitelných sémanticky, gramaticky, syntagmaticky, pragmaticky aj., pokusit se pro tyto třídy nalézt ideální způsob lexikografického zpracování a ten následně aplikovat na vznikající lotyšsko-český slovník (vůbec první příručku svého druhu). K tomuto účelu bylo využito moderních nástrojů, jež v nedávné minulosti radikálně změnily způsob lexikografické práce: jednak jazykových korpusů, reprezentujících dnes autentický jazykový úzus, jednak specializovaného lexikografického softwaru Tshwane-Lex, ve kterém je budována lexikální databáze lotyštiny, z níž bude následně sestaven samotný slovník. Tradiční zdroje nebylo možno zcela eliminovat, už kvůli omezené velikosti lotyšského korpusu, a autor práce byl tak nucen scelit tradiční a moderní lexikografické metody. Primárním zdrojem pro něj však zůstávají korpusová data, ta jsou konfrontována s popisem v aktuálních lexikografických i obecně jazykovědných příručkách a případné rozpory jsou řešeny ve prospěch údajů nabytých z korpusů. Lexikografická řešení jsou v této práci průběžně prezentována prostřednictvím vzorových hesel. V mnoha případech jde o řešení inovační, dosud nevyužitá v žádném z dvoujazyčných slovníků s lotyštinou jako zdrojovým jazykem: i to je důsledek změny přístupu a obměny materiálové báze. Kromě toho byly revidovány a zpřesněny rovněž některé poznatky čerpané z aktuální lotyšské mluvnice.

Ve studii je představena slovníkářská práce v celé své komplexnosti. Chystaný slovník je nejprve vymezen z hlediska typologického, stanoveny jsou jeho cíle a adresát. Následně jsou rozvrstveny zdroje pro materiálovou základnu a představen technický aspekt tvorby slovníku. Dále se postupuje od sestavení hesláře přes tvorbu jednotlivých heslových statí a doprovodného informačního aparátu až po finální fázi editace a kontrolu rukopisu slovníku. Využívá se přitom kombinované teoreticko-praktické metody: jednotlivé problémy jsou teoreticky nahlíženy s oporou v autoritativních jazykovědných příručkách a konfrontovány jak s korpusovými daty, tak s popisem ve slovnících, poté jsou popsány rozhodovací strategie vedoucí k předkládaným lexikografickým řešením. Práci uzavírá trojice případových studií, kde je na vybraných lexémech podána lexikografická praxe z bližší perspektivy.

Práce představuje záruku profesionálního, nediletantského přístupu a potřebného teoretického rámce k vlastní lexikografické činnosti a je tak pouhým předstupněm k většímu úkolu: vytvoření lotyšsko-českého slovníku. V tomto smyslu zůstává otevřená a do značné míry podmíněna jeho zdárným vydáním.

**Klíčová slova:** dvoujazyčná lexikografie, kontrastivní lexikografie, korpusová lexikografie, lotyšsko-český slovník, lexikografická databáze, Tshwane-Lex

## **Abstract**

**Title:** Comparative aspects of Latvian and Czech lexicons: Materials for assembling a Latvian-Czech dictionary

**Autor:** Mgr. Michal Škrabal

**Department:** Institute of the Czech National Corpus

**Supervisor:** prof. PhDr. František Čermák, DrSc.

The primary aim of this work is to classify the Latvian lexicon, or better its relevant segment, into individual groups, definable semantically, grammatically, syntagmatically, pragmatically, and so on, and to attempt to find for these classifications an ideal method of lexicographical adaptation and apply it to an emerging Latvian-Czech dictionary (the very first manual of its type). To this end, modern instruments were utilized which, in the recent past, have radically altered the methodology of lexicographical work: on the one hand, the linguistic corpora, which nowadays represent authentic, linguistic usage and, on the other hand, the specialized lexicographic software TshwaneLex, in which a lexical database of Latvian is constructed and from which the dictionary itself will be subsequently constructed. Because of the limited size of the Latvian corpus it was not possible to completely eliminate traditional sources, and the author of the work was forced to consolidate traditional and modern lexicographical methods. His primary source however remained the corpus data, which are directly compared with descriptions in current lexicographical and general linguistic manuals and potential variances are resolved in the results of the data acquired from the corpus. The lexicographical solutions in this work are concurrently presented by means of model entries. In many cases, it concerns an innovative solution, to this time not utilized in any bilingual dictionary having Latvian as a source language: it is also the result of changes of the approach and



alterations of the material base. In addition, some findings from contemporary Latvian grammar were checked and made more precise.

In the study, lexicographical work is presented in all of its complexity. The forthcoming dictionary is first demarcated from a typological point-of-view, its goals and intended readers defined. Subsequently, the sources for the material base are stratified and the technical aspect of the dictionary's creation is presented. Further on, it proceeds from the assembling of a list of entries to the formation of individual lead entries and accompanying information apparatus all the way up to the final phase of editing and inspection of the manuscript of the dictionary. At the same time, it makes use of a combined theoretical-practical method: individual problems are consulted theoretically with backing in authoritative, linguistic manuals and compared both with corpus data and with their description in dictionaries. Finally the decision strategies are described, leading to the presented lexicographic solution. A trio of case studies closes the study, where lexicographic practice from a closer perspective is applied to selected lexemes.

The work constitutes a guaranty of a professional, non-dilettantish approach and a necessary theoretical framework to a separate lexicographic activity. It is, then, merely a step towards a greater task: the creation of a Latvian-Czech dictionary. In this sense, this work remains open and is to a certain extent conditioned by the successful publication of the dictionary.

**Keywords:** bilingual lexicography, contrastive lexicography, corpus lexicography, Latvian-Czech dictionary, lexicographical database, TshwaneLex

## Kopsavilkums

**Darba nosaukums:** Latviešu leksika no divvalodu leksikogrāfijas viedokļa: Teorētiskais pamats latviešu-čehu vārdnīcas izveidošanai

**Autors:** Mgr. Mihals Škrabals (*Michal Škrabal*)

**Katedra:** Čehu nacionālā korpusa institūts (*Ústav Českého národního korpusu*)

**Disertācijas vadītājs:** prof. PhDr., zinātņu doktors Františeks Čermāks (*František Čermák*)

Darba primārais mērķis ir sadalīt latviešu leksiku, respektīvi, tās nozīmīgo segmentu, atsevišķās grupās, kuras definējamas semantiski, gramatiski, sintagmatiski, pragmatiski u. c., mēģināt atrast šīm klasēm ideālo leksikogrāfiskās apstrādes paņēmieni un to pēc tam lietot jaunveidojamai latviešu-čehu vārdnīcai (kas būs vispār pati pirmā šāda veida rokasgrāmata). Šim nolūkam izmantoti modernie instrumenti, kuri nesenā pagātnē radikāli mainījuši leksikogrāfijas darba veidu: tas attiecas gan uz valodas korpusiem, kuri reprezentē šobrīd autentiskās valodas normas, gan uz specializēto leksikogrāfisko programmatūru *TshwaneLex*, kurā tiek veidota latviešu valodas leksikas datu bāze, no kuras pēc tam tiks veidota pati vārdnīca. Pilnībā nebija iespējams izslēgt arī tradicionālos resursus, jau tāpēc vien, ka latviešu korpusi ir ierobežoti, tāpēc darba autors bija spiests apvienot tradicionālās un modernās leksikogrāfijas metodes. Tomēr par autora primāro resursu izmantoti korpusu dati, konfrontējot tos ar aprakstiem aktuālajās leksikogrāfijas un vispārējās valodniecības rokasgrāmatās, un iespējamās pretrunas risinātas par labu datiem, kuri iegūti no korpusiem. Leksikogrāfijas risinājumi šinī darbā sistemātiski prezentēti ar parauga šķirķļu starpniecību. Daudzos gadījumos tie ir inovatīvi risinājumi, kuri līdz šim nav izmantoti nevienā no divvalodu vārdnīcām, kurās latviešu valoda ir avota valoda: arī tas ir pie-

ejas maiņas un materiālu bāzes nomaiņas rezultāts. Vienlaikus revidētas un precizētas dažas atziņas, kuras iegūtas no aktuālās latviešu gramatikas.

Pētījumā parādīts vārdnīcas sastādītāja darbs visā savā daudzveidībā. Gatavojamā vārdnīca vispirms determinēta no tipoloģijas viedokļa, noteikti tās mērķi un adresāts. Pēc tam pa slāņiem izvērsti resursi materiāla bāzei un parādīts vārdnīcas veidošanas tehniskais aspekts. Tālāk darbība noris, sākot ar šķirklju saraksta sastādīšanu, turpinot ar atsevišķo šķirklju apraksta un pavadošā informatīvā materiāla veidošanu un beidzot ar noslēdzošo vārdnīcas rokraksta rediģēšanas un pārbaudes fāzi. To darot, izmantota teorētiski praktiskā metode: atsevišķās problēmas teorētiski izvērtētas, balstoties uz autoritatīvām valodas zinātnes rokasgrāmatām un konfrontētas kā ar korpusu datiem, tā ar aprakstu vārdnīcās, un pēc tam aprakstītas izlemjošās stratēģijas, kuras noved pie uzrādītajiem leksikogrāfiskajiem risinājumiem. Darbu noslēdz trīs atsevišķo gadījumu pētījumi, kuros, izmantojot izvēlētas leksēmas, no tuvākas perspektīvas demonstrēta leksikogrāfijas prakse.

Darbs sniedz garantiju tam, ka pieeja pašai leksikogrāfiskajai darbībai ir profesionāla un bez diletantisma, veidojot tai nepieciešamo teorētisko ietvaru, un tādējādi ir tikai priekšposms lielākam uzdevumam: latviešu-čehu vārdnīcas izveidošanai. Šādā nozīmē tas paliek atvērts un ievērojamā mērā saistīts ar sekmiņu vārdnīcas izdošanu.

**Atslēgvārdi:** divvalodu leksikogrāfija, kontrastīvā leksikogrāfija, korpusu leksikogrāfija, latviešu-čehu vārdnīca, leksikogrāfijas datubāze, TshwaneLex

## Obsah

<b>Část první.....</b>	<b>14</b>
<b>Úvod.....</b>	<b>14</b>
1.1 Cíl práce .....	14
1.2 Stručný nástin práce .....	16
1.3 Metodologie a postup při zpracování.....	17
<b>Část druhá.....</b>	<b>21</b>
<b>Východiska pro tvorbu LČ.....</b>	<b>21</b>
2.1 Situace lotyšská a česká .....	21
2.2 LČ z hlediska typologického, jeho adresát a cíle .....	27
2.3 Lexikografické zdroje pro tvorbu LČ .....	30
2.3.1 Korpusy .....	31
2.3.2 Slovníky .....	37
2.3.3 Další lingvistické zdroje .....	42
2.3.4 Internetové zdroje .....	42
2.3.5 Excerpta .....	43
2.4 Technický aspekt tvorby LČ .....	45
<b>Část třetí.....</b>	<b>62</b>
<b>Teoretický rámec a metodika zpracování LČ .....</b>	<b>62</b>
3.1 Možnosti makrostruktury .....	62
3.1.1 Heslář LČ .....	62
3.1.2 Typy hesel v LČ.....	65
3.2 Možnosti mikrostruktury .....	75
3.2.1 Lemma .....	76
3.2.2 Informace fonologická.....	77
3.2.3 Informace gramatická .....	82
3.2.3.1 Gramatická charakteristika substantiv v LČ.....	85
3.2.3.2 Gramatická charakteristika verb v LČ.....	101

3.2.3.3 Gramatická charakteristika konjunkcí v LČ.....	116
3.2.4 Informace sémantická .....	118
3.2.5 Informace syntagmatická .....	127
3.2.6 Informace diasystematická .....	131
3.2.7 Informace uzuální .....	138
3.2.8 Informace referenční .....	140
3.3 Možnosti megastruktury .....	141
3.4 Editace a revize rukopisu LČ .....	142
<b>Část čtvrtá.....</b>	<b>144</b>
<b>Vybrané otázky dvoujazyčné lexikografie a jejich řešení v LČ ....</b>	<b>144</b>
4.1 Zpracování polysémního substantiva <i>biedrs</i> v LČ .....	145
4.2 Varianty konjunkcí <i>turpretī</i> a <i>iekams</i> a jejich zpracování v LČ ....	152
4.3 Zpracování homonymního lexému <i>ārs</i> v LČ.....	156
<b>Část pátá.....</b>	<b>164</b>
<b>Závěr.....</b>	<b>164</b>
<b>Bibliografie.....</b>	<b>168</b>
<b>Seznam obrázků a tabulek .....</b>	<b>Chyba! Záložka není definována.</b>

# Část první

## Úvod

### 1.1 Cíl práce

Objekt našeho výzkumu i perspektiva, jíž na něj nahlížíme, vyjadřuje titul předkládané práce, cíl výzkumu pak její podtitul. V centru naší pozornosti stojí lotyšský lexikon, nazíraný prizmatem dvoujazyčné, konkrétně lotyšsko-české lexikografie; intencí je úměrně relevantní část<sup>1</sup> tohoto lexikonu analyzovat a její popis vtělit do chystaného lotyšsko-českého překladového slovníku (dále v textu: LČ). Vzhledem k imanentní heterogenitě slovní zásoby každého přirozeného jazyka je třeba lotyšský lexikon účelně rozdělit do stejnorodých tříd a ty charakterizovat zvlášť. Je nabíledni, že přístupy k takovému postupu mohou být různé, stejně jako výběr problémů, jež je třeba popsát. Naše volba případů k analýze je motivována především postavením konkrétních jevů v systému: nejde nám o kontrastivní rozbor celého lotyšského lexikonu, ale zajímají nás primárně úkazy centrální, jež budou v LČ hojně zastoupeny. Pro účely této práce by i omezená selekce měla být dostačující. Nezastíráme tu ani skutečnost, že některé specifické složky lexikonu (frazologie, terminologie, slang aj.) byly vyděleny od části bezpříznakové a jejich zpracování odsunuto na později, jistá zde předkládaná řešení jsou proto pouze pracovní a budou podrobena další úpravě.

---

<sup>1</sup> Jde o minimálně dvoustupňovou redukci výchozího materiálu, svým způsobem nevyčerpatelného. První nezbytné zjednodušení tkví v samotné snaze o popis v příručce lexikografického typu: slovník, ať sebevětší, zůstane vždy jen pouhou aproximací jazyka, zvláště pak jeho parolové stránky. Druhé omezení vychází z povahy LČ, jenž je plánován jako slovník středního rozsahu, což oproti jiným, ambicióznějším projektům vyžaduje střídmejší perspektivu.

Práce by měla posloužit jako nezbytný teoretický fundament pro zamýšlené lexikografické dílo a zajistit tak mimo jiné profesionální, nediletantský přístup k slovníkářské práci, (nejen) v dnešní době hojně rozšířený.<sup>2</sup> Chystaný lotyšsko-český slovník je vůbec první – a nejspíš nadlouho i poslední – příručkou svého druhu, a měl by tudíž být co možná nejobsáhlejší, respektive co nejvšestrannějším nástrojem, pomocí něhož lze přeložit širokou škálu textů: od klasického literárního díla přes aktuální publicistiku až po odborný text. Cílem rozhodně není zaplnit stávající mezeru narychlo stvořeným, okleštěným kompilátem existujících slovníků, vycházejícím vstříc pouze nejzákladnějším potřebám uživatelů. Toto vědomí odpovědnosti klade na autora slovníku nemalé nároky: k práci je třeba přistupovat maximálně zodpovědně, nejen s dobrou praktickou znalostí výchozího jazyka, ale i s nezbytnou teoretickou průpravou a v neposlední řadě i s akribií.

Výběr tématu je dán do značné míry pragmatickými pohnutkami,<sup>3</sup> a sice absencí jakéhokoliv lotyšsko-českého slovníku. Tento dlouhodobý deficit negativně pocítují všichni čeští zájemci o lotyštinu a také lotyšští zájemci o češtinu – a přestože jde o skupiny relativně nepočetné, samotný fakt, že jsou na prostřednictví třetího jazyka odkázáni obyvatelé stejného soustátí, je znepokojivý. Jde tu však též o vitalitu celé české letonistiky, která bude bez této základní oborové příručky vždy o krok pozadu.

---

<sup>2</sup> Navzdory četným rádobý lexikografickým internetovým projektům, ale též seriózněji míněným tištěným slovníkům, kdy aprobaci autorovi poskytne několikaletý post velvyslance v cizí zemi, či dokonce čirý entuziasmus, pevně věříme v platnost výroku Franze Josefa Hausmanna: „Zweisprachige Wörterbücher brauchen die Universität.“ (citováno podle Vachková 2007: 3)

<sup>3</sup> Autor práce by se nicméně rád přiznal – takto pokoutně v poznámce pod čarou – i k čistě osobní motivaci. Vzhledem k absenci jakéhokoliv předchůdce v podobě lotyšsko-českého slovníku (byť zastaralého; existují jen dva česko-lotyšské protějšky) jde o počín průkopnický. A pionýrství vždy bývalo lákavé a vzrušující. Autor se vydal na neprošlapanou cestu lexikálním vesmírem, bohdá se v něm neztratí...

## 1.2 Stručný nástin práce

Tato studie je členěna do pěti částí. V úvodu jsou vyloženy její cíle, je zdůvodněn výběr tématu a podán přehled o jeho dosavadním odborném zpracování a nakonec je vyložena i metodologie, pomocí níž bude studie vypracována.

Ve druhé části obhlíží autor výchozí situaci, v níž se s projektem LČ ocitá. Analyzovat je třeba v první řadě aktuální stav lotyšské jazykovědy (kap. 2.1), zejména v relevantních disciplínách, jimiž jsou lexikografie, kvantitativní a korpusová lingvistika. Shrnuty jsou též dosavadní nečetné výtěžky česko-baltské lexikografie. Dále je zapotřebí rozhodnout, komu bude LČ určen – od adresáta slovníku (2.2) se obecně odvíjí všechny další postupy, popsány v následující kapitole. Takto je LČ vymezen z hlediska typologického, stanoveny jsou jeho cíle. Následně jsou rozvrstveny zdroje pro materiálovou základnu (2.3): jazykové korpusy, slovníky a jazykovědné příručky a doplnkově rovněž internetové prameny a excerpta. V posledním oddíle (2.4) je pozornost věnována technickému aspektu tvorby LČ: popisu specializovaného lexikografického softwaru TshwaneLex, ve kterém vzniká lexikální databáze, sloužící jako podklad pro samotný slovník (případně i pro další projekty).

Třetí část, tvořící jádro celé studie, obsahově pokrývá jednotlivé fáze vlastní lexikografické práce: od sestavení hesláře (3.1) přes tvorbu jednotlivých heslových statí (3.2) a doprovodného informačního aparátu (3.3) až po finální fázi editace a kontrolu rukopisu slovníku (3.4). Po stránce metodologické je teoreticko-praktická: jednotlivé problémy jsou teoreticky nahlíženy s oporou v autoritativních jazykovědných příručkách a konfrontovány jak s korpusovými daty, tak s popisem ve slovnících, načež jsou popsány rozhodovací strategie vedoucí k předkládaným lexikografickým řešením.

Čtvrtá část ukazuje na třech případových studiích slovníkářskou práci v její komplexnosti. Lexikograf má na začátku pouze lemma, které musí doplnit o zbytek heslové statí; od analýzy, vycházející z korpusových dat a popisu



lexému ve slovníkových příručkách, gramatikách a jiných zdrojích, se dobírá k syntéze. Obvykle přitom řeší vícero lexikografických problémů současně.

Konečně, v závěrečné, páté části je podáno shrnutí dosažených poznatků, práce je následně uzavřena soupisem použité literatury a seznam tabulek a obrázků.

### **1.3 Metodologie a postup při zpracování**

Předkládaná studie je metalexikografická, věnujíc se teoretickému aspektu lexikografie; zároveň má posloužit jako východisko pro tvorbu dvoujazyčného slovníku, což je úkol ryze aplikovaný, představující praktický pól lexikografie.<sup>4</sup> Tato dvojdomost se přirozeně odráží i v kombinaci přístupu teoretického a praktického, analytického a syntetického.

Z jiných metalexikografických textů čerpáme zejména inspiraci: širě nabízených řešení je ohromná; je tudíž prakticky nemožné přijít se skutečně originální ideou. Ostatně není to ani naší ambicí: lexikografický popis by měl být v první řadě co nejvěrnější skutečnému jazykovému úzu, tedy nevycházet z jakéhosi lexikografova nereálného konstruktu, jemuž by se podřizovala autentická data. Neméně důležitý je při zpracování však také ohled na uživatele slovníku: informace musí být pohodlně a rychle vyhledatelná a snadno interpretovatelná. U dvoujazyčného slovníku jde především o to zachovat v nejvyšší možné míře význam a (textovou) funkci výrazu z výchozího jazyka, jinými slovy: snažit se o co nejširší substituovatelnost daného prvku mezi maximem potenciálních textů, kde se vyskytuje, v obou jazycích (viz též 2.2).

Zvláště užitečné shledáváme prakticky zaměřené lexikografické manuály z nedávné doby (Svensén 2009, Atkins–Rundell 2008, van Sterkenburg 2003), jež jsou vítaným doplňkem monografií starších (Zgusta 1971, Béjoint 2000,

---

<sup>4</sup> Diskusi kolem vymezení obsahu lexikografie, ale též jejího postavení vůči jiným lingvistickým odvětvím (namně lexikologie) shrnuje např. Sekaninová (1993: 20–25) či Svensén (2009: 2–5)

Landau 2001, Hartmann 2001) či specializovaných (Bergenholtz–Tarp 1995) v čele s monumentálním oborovým kompendiem, vydaným ve třech svazcích v letech 1989–91 (Hausmann et al. 1989–1991). Tyto zdroje doplňujeme o nepříliš četné české (včetně slovenských) prameny (kromě dílčích studií uceleně např. Sekaninová 1993, Čermák–Blatná 1995, Vachková 2007, 2011) a ještě kusejší lotyšské (Baldunciks 2012, Andronovs 1998, k metalexikografii česko-lotyšské ojedinělé též Nikuļceva 2013).

Metodologicky se opíráme o koncept kontrastivní lexikografie, potažmo lexikologie (mj. Kopeckij 1974; Bezděk 1974; Filipec 1985; Sekaninová 1993; Jarošová 2000, 2008a, 2008b; Vachková 2007 aj.), s využitím možností, jež nabízejí dnešnímu slovníkáři jazykové korpusy (k možnostem, ale též mezím korpusové lingvistiky viz Cvrček–Kováříková 2011). Korpusová data způsobila v lexikografii převratné (a nenávratné) změny:<sup>5</sup> dochází k přehodnocování např. poměru syntagmatiky a paradigmaticky ve prospěch prvně jmenované (Čermák 2010b: 565), vztahu centra a periferie, preskripce vs. deskripce, vnímání jazykovědcovy introspekce (Vachková 2007: 32) aj., dokonce i k hledání nového přístupu k jazyku obecně.<sup>6</sup>

Náš přístup ve vztahu ke korpusovým metodám má blíže k pólu *corpus-based* (srov. však možné přechody k *corpus-driven* přístupu např. v kap. 4.3), což je dáno mimo jiné i omezeným množstvím dat, jež by nedostačovala k vytvoření lotyšsko-českého slovníku v předpokládaném rozsahu a kvalitě. Autor studie tuto svou pozici na pomezí starého a nového způsobu lexikografické práce dříve obrazně nazval „mostem“ mezi tradičním a moderním přístupem k lexikografii (Škrabal 2011),<sup>7</sup> aniž by jej považoval apriori za handicap. Od

---

<sup>5</sup> Často se hovoří o lexikografické revoluci – viz např. Svensén 2009: 45–47.

<sup>6</sup> „[...] collocation becomes the absolute keystone of the study of lexis and semantics, and thus (due to the centrality of meaning in Firthian linguistic theory) the keystone of the entire approach to language.“ (McEnery–Hardie 2012: 132).

<sup>7</sup> „Ve své povaze pak zvláště tradiční manuální lexikografie doznává při svém přerodu zásadní změnu především ve formě a technice práce (je rychlejší, komplexnější, exaktnější a důslednější), danou jinou kvalitou materiálového zdroje (je bohatší, odstíněnější, spolehlivější a věrnější); ani cíl lexikografie nezůstává ovšem plně stejný, protože nový kontext práce s sebou nese řadu nových možností. Obecně však jde při tomto přechodu o plnou návaznost na všechny využitelné mentální procesy a procedury lexikografie tradiční.“ (Čermák 1995b: 52)

stávajícího popisu lotyšské gramatiky a lexikonu nechceme abstrahovat jen proto, že stojí na „předkorpusových“ základech.

Vycházíme tudíž ze současné jazykové normy lotyštiny, reprezentované novou lotyšskou gramatikou,<sup>8</sup> která je oproti své předchůdkyni z přelomu 50. a 60. let minulého století (MLLVG I, II) mnohem méně rigidní a preskriptivní a připouští různé možné přístupy k jednotlivým jazykovým jevům i jejich častou nevyhraněnou či hraniční povahu, prohlašujíc se za mluvnicki synchronně deskriptivní (LVG: 19), „podmíněně normativního charakteru“ (*nosacīti normatīvs raksturs*).<sup>9</sup> Korpusy (případně i další data) nám pak slouží jako materiálová základna pro hledání odchylek od této normy. Tyto diskrepance zaznamenáváme do databáze jakožto signál potenciálních změn v paradigmatu, jinými slovy: jakožto příznak (nevyhnutelného) jazykového vývoje.<sup>10</sup> Odchyly od normy vyskytující se neojediněle, neřkuli pravidelně, by měl LČ postihnout, a to i navzdory aktuálnímu lexikografickému popisu v lotyšských jednojazyčných i překladových slovnících. V našem chápání činí z lexému lexikografickou entitu primárně jeho rekurentní charakter a případný nesoulad se stávající normou je možno naznačit vhodně zvoleným kompenzačním prostředkem, nelze však odmítnout zahrnutí lexému do LČ na základě jeho „diskvalifikujícího“ markeru ve výkladovém slovníku (viz též 3.2.6).

Teoretické vývody pak transformujeme do podoby četných ukázkových hesel, jež jsou aktuálním výstupem z naší lexikální databáze (tudíž mají leckdy pouze pracovní charakter a mělo by na ně být také tak nahlíženo); často heslové statě nepřetiskujeme celé, ale jen jejich relevantní část odpovídající právě vykládané problematice.

---

<sup>8</sup> *Latviešu valodas gramatika* z r. 2013, resp. její 2. vydání z r. 2015 (viz též dále 2.3.2.1), srov. též koncepci této gramatiky podanou v LVGK (2008).

<sup>9</sup> „[...] stanovení normy není primárním cílem nebo hlavním úkolem. S oporou v tradici určování gramatických norem jsou v nové gramatice lotyštiny naznačeny tendence ke změnám či vývoji norem.“ (LVGK: 14–15)

<sup>10</sup> Nic se v tomto ohledu nezměnilo na výroku L. Kopeckého, proneseném již v roce 1972: „Jednotky slovní zásoby se musí inventarizovat v tom použití, které je doloženo texty vybranými pro excerpci. Možné neshody s doporučeními přijaté normy může slovník označovat, nemůže to však být překážkou pro inventarizaci.“

Anglicko- a německojazyčné citáty spolu se slovenskými ponecháváme v originále, citáty lotyšské převádíme do češtiny. Cizí příjmení autorek s výjimkou odkazů a bibliografie soustavně přechylujeme.

Poznámka k lotyšské grafematice: vodorovná čárka nad písmenem označuje délku samohlásky, grafémy *ā, ē, ī, ū* tak odpovídají příslušným dlouhým vokálům v češtině. Měkké souhlásky *ģ, ķ, ņ* korespondují s českými grafémy *d', t', ň*; souhláska *ļ* svůj český protějšek nemá, označuje měkkou variantu hlásky /l/.

## Část druhá

### Východiska pro tvorbu LČ

#### 2.1 Situace lotyšská a česká

2.1.1 Lotyšská lexikografie může čerpat z dlouhodobé tradice, sahající až do první poloviny 17. století, byť celých prvních 250 let nebyli autory ani adresáti slovníků většinou etničtí Lotyši (Jansone–Veisbergs 2013: 435). První překladový slovník, německo-lotyšský *Lettus* z roku 1638,<sup>11</sup> byl napsán německým pastorem Georgem Manzelem a stál u zrodu spisovné lotyštiny (Štoll 2013: 40–41). Navzdory této bohaté tradici i skutečnosti, že v lotyšském prostředí, podobně jako u většiny malých národů, dominují slovníky překladové nad výkladovými (Jansone–Veisbergs 2013: 435), je počet jazykových kombinací lotyšských dvoujazyčných slovníků překvapivě malý: nepřevyšuje patnáct (ibid.: 441–443, 449–451). Ze slovanských jazyků, které by nejbližší odpovídaly naší výchozí situaci, a mohly být tudíž využity jako zdroje, tvoří pravou stranu lotyšských překladových slovníků ruština a polština, oba příslušné slovníky využíváme (viz dále 2.3.2.2).

Aktuální situaci na poli lotyšské bilingvní lexikografie dobře vystihuje citát J. Baldunčikse:

„Přestože lotyšsko-anglické a anglicko-lotyšské slovníky jsou v reálné současné jazykové situaci nejdůležitější, [...Andrejs Veisbergs ve svém výzkumu konstatuje, že...]

---

<sup>11</sup> Heslář čítající přibližně 7000 hesel je řazen abecedně. Autor užívá diakritiku k zachycení specifík lotyšského fonologického systému, slovník je vytištěn frakturou. Lemma je od lotyšského protějšku odděleno lomítkem, za zmínku stojí pokus o zachycení diasystematické informace (explicitní označení germanismů, teritoriální příslušnost některých lotyšských ekvivalentů) či přítomnost víceslovných lemmat (ustálených spojení nefrazeologické povahy). Ke slovníku je připojena tematicky uspořádaná část *Phraseologia Lettica*, kombinující přístup onomaziologicko-sémaziologický (je členěna do 51 oddílů, každý z nich je řazen abecedně). (Jansone–Veisbergs 2013: 435). Slovník byl digitalizován a jako součást Korpusu starých lotyšských textů je k dispozici online, včetně frekvenční statistiky, na

[http://www.korpuss.lv/senie/source.jsp?codificator=Manc1638\\_L](http://www.korpuss.lv/senie/source.jsp?codificator=Manc1638_L).

jejich tvorba (včetně těch elektronických) zůstala na úrovni starých metod. Zásadní obtíže nepředstavuje vytváření anglické části nebo hledání ekvivalence, jakkoliv jde o časově náročné [úkoly]. Základní problém se skrývá v lotyšské části *per se* [...]. V rozvoji nejvíce brání to, že chybí korpus lotyštiny, tudíž lotyšští lexikografové nemohou objektivně vytvořit lotyšskou část: jasno není dokonce ani v otázce aktualizace hesláře, v použitelnosti, frekvenci a významu nově přidanych slov a slovních spojení. [...] Kvůli chybějícímu korpusu nebo alespoň rozsáhlého a moderního lotyšského hesláře se při doplňování slovníků nezdá odhalují staré lexikální mezery: slova zapomenutá, „proklouznuvší“. Absence obsáhlého a vyváženého korpusu dnešního lotyšského lexika je závažnou překážkou rozvoje všech dvoujazyčných slovníků s lotyštinou jako zdrojovým jazykem.“ (Baldunčiks 2012: 11–12, srov. Baldunčiks 2015)

Citelně tu tedy chybí dostatečně veliký reprezentativní korpus současné lotyštiny, byť první kroky byly v tomto směru již podniknuty (blíže v kap. 2.4). Od toho se odvíjí i relativně slabá pozice korpusové lingvistiky v Lotyšsku, jejíž potenciál tu dosud nebyl plně rozvinut, na rozdíl třeba od lingvistiky počítačové. Stávající synchronní korpus lotyštiny slouží jako materiálová základna pro další lingvistický výzkum, především v oblasti sémantiky a syntaxe (viz např. série článků Gunty Nešporeové věnující se nejrůznějším aspektům pohybových sloves – Nešpore 2009 aj.; další korpusově založené studie zmiňují Andronova–Andronovs 2011: 48), avšak většinový postoj lingvistické obce ke korpusovým datům zůstává spíše rezervovaný.<sup>12</sup>

2.1.2 Kvantitativní výzkum lotyštiny probíhá v Lotyšsku od poloviny 60. let minulého století, byť zčásti dovršen vydáním čtyřsvazkového Frekvenčního slovníku lotyšského jazyka (*Latviešu valodas biežuma vārdnīca*, LVBV) v letech 1966–1976. Na vzorku 892 tisíc slov<sup>13</sup> excerpovaných z 892 zdrojů

---

<sup>12</sup> Srov. však vloni vydaný lotyšsko-estonský slovník (LE) spolu s jeho protějškem, oba o velikosti cca 40 tisíc hesel. Vznikly v enormně krátké době dvou let na základě paralelního korpusu, jehož výsledky byly autory slovníku zrevidovány a doplněny. Slovník je dostupný též online na <http://www.ee-lv.lv/lv/vardnica>.

<sup>13</sup> Srov. obdobný projekt český (Jelínek – Bečka – Těšitelová 1961).

(z každé sondy byl vybrán ucelený fragment o rozsahu tisíc slov) ve třech různých textových typech – publicistice, beletrii a vědecko-technické literatuře – vznikl index celkem 31 039 lexémů, formou frekvenčního seznamu jsou podány ty s minimální frekvencí 3 (úhrnem 15 979 jednotek). Na LVBV navázali další badatelé, sledující většinou praktické aspekty zkoumání frekvenční špičky a jejich využití pro pedagogické účely (přehledy lexikálního minima, např. Kuzina 1998a).

Dnes lze kvantitativní vlastnosti lotyštiny pohodlně zkoumat pomocí jazykových korpusů, čímž se velmi zjednodušila práce zejména lexikografům: dostupné údaje o frekvenční špičce tak pozitivně ovlivní rozhodování o složení hesláře (3.1.1). LVBV může nicméně posloužit k ověření frekvenčních údajů či ke sledování vývojových tendencí.

2.1.3 V českých zemích existoval zájem o baltskou lexikografii už minimálně od první poloviny 19. století, omezoval se však výhradně na oblast akademickou;<sup>14</sup> z hlediska ryze praktického zůstávala oblast česko-baltské lexikografie netknuta až do konce 80. let 20. století: ač jazyky geneticky (a relativně i regionálně) blízké a kulturně spjaté,<sup>15</sup> nevynutily si dlouhou dobu žádnou utilitární příručku slovníkového typu. Za určitou výjimku lze považovat útlou publikaci Československo-lotyšské společnosti *Slovníček česko-lotyšský* z roku 1928, obsahující přibližně 3000 hesel (obr. 1). Dojem z tohoto průkopnického díla kazí mnohde chybějící lotyšská diakritika i nepřesnosti ve stručném gramatickém úvodu. Autor spisku Hanuš Entner, profesí železniční úředník,<sup>16</sup> stěží mohl vytvořit slovník s vědeckými ambicemi; ostatně sám to nepřímou přiznává

---

<sup>14</sup> Pozornost badatelů přitahovaly především vzájemné paralely, zděděné z období balto-slovanské jednoty, v čele se společným lexikálním fondem, podepřeným etymologickými výklady. Takto je koncipován Čelakovského *Litevský slovník* (litevština tu zastřešuje všechny tři baltské jazyky, tedy i lotyštinu se starou pruštinou), nedokončený *Etymologický slovník litevštiny* Josefa Zubatého či *Baltisch-Slavisches Wörterbuch* Reinholda Trautmanna, vznikající na půdě pražské univerzity a vydaný v Göttingenu v roce 1923.

<sup>15</sup> Podrobněji k tradici česko-lotyšských kulturních vztahů, pokrývajících posledních pět století, Štoll 2013.

<sup>16</sup> Blíže k jeho osobnosti ve vztahu k Lotyšsku Malý 2013 či Miškovský 2013.

v lapidárním doslovu: „V češtině nebylo dosud lotyšského slovníku, jest tedy tato práce první svého druhu. Jako taková budiž přijata.“ (Entner 1928: 63)

budoucí nākams, turpmaks budoucnost, nākotne budova, būve, ēka Bůh, Dievs býk, bulis, vērsis	bydlet, mājot byt, dzīvoklis býti, būt bylost, būte
<b>C</b>	
car, cars celnice, muitnica celník, muitnieks celý, viss cena, maksa, vērtība ceník, cenu saraksts cep, sprigulis cesta, ceļš cestování, ceļojums cestovat, ceļot cestující, ceļotājs cibule, sīpols ciferník, ciparnica cihla, ķieģelis cíl, mērķis cín, alva císař, ķeizars cit, sajūta cukr, cukurs	cítiti, sajust citron, citrons cizí, svēšs cizina, ārzeme svešums, cizinec, ārzemnieks, sveš- nieks clo, muiža co, ko, kas, ka cop, bize ctihodný, cienīgs ctitel, cienītājs ctíti, cienīt ctīzādost, godkāriba tikums, ctnost cukrār, konditors cukrārštvi, konditoreja cvičení, vingrinājums cvičiti, vingrināt cvrček, circeņis
<b>Č</b>	
čaj, tēja čalouník, tapsetājs čáp, gandrš čapka, cepure čára, līnija	čaroděj, burvis čarovný, burvīgs čas, laiks časopis, laikraksts část, daļa
2*	19

**Obrázek č. 1:** Ukázka ze *Slovníčku česko-lotyšského* H. Entnera (1928)

Po anexi pobaltských republik roku 1940 na sebe zprostředkovatelskou roli vzala sovětská *lingua franca*, ruština. Právě o ni (a zčásti též o němčinu) se opíral autorský kolektiv, pracující na *Česko-lotyšském slovníku* (obr. 2, dále v textu: ČL); ten vyšel v roce 1988 a čítal okolo 22 tisíc hesel. Jeho následníkem, využívajícím však už do značné míry nových lexikografických metod, především jazykových korpusů (v tomto případě korpusu synchronní češtiny SYN2000), je stejnojmenný slovník Sandry Nikuļcevové z roku 2006 (obr. 2, dále v textu: ČLN). Ač přibližně stejného rozsahu, je na první pohled mnohem obsáhlejší, což je dáno daleko větším množstvím kolokací a frazémů i podrobnějším a vylepšeným poznámkovým aparátem.



etnický etniský	etnický [-ny-] adj <i>etnický</i> s; ~á skupina/ menšina <i>etnická grupa/ minoritáta</i> ; ~é konflikty/ čísky <i>etnické konflikty/ etnická tířšana</i>
etnografie s. etnografija	etnikum [-ny-] n, (-a) <i>etnos</i> s m
evakuace s. evakuacija	étos m, (-u, L -u) <kníž.> <i>étos</i> s m
evakuační evakuācijas-; ~ stře-	etudja [-tý-] f, (-y) <hud.> <i>etúde</i> f, (Gpl -zu)
disko — evakuācijas punkts	EU [é-ú] zkr. (Evropská unie) <i>ES</i> [é-es] ( <i>Eiropas Savienība</i> )
evakuovat p., i. evakuēt	euforie f, (-e) <i>eifonija</i> f
evangelický rel. evaņģēliskš; evaņ-	eurjo n, (-a, L -u) <i>eiro</i> m, (neskl.)
ģēlija-	evakuace f, (-e) <i>evakuacija</i> f
evangelium n. rel. evaņģēlijs	evakuovat (im)pf, (-uje) <i>evakuēt</i> (II.); (podrobit se evakuaci) <i>evakuēties</i> (II.)
eventualita s. iespējamība; iespē-	evangelický adj 1 (týkajci se evangēlia) <i>evaņģēlija</i> (G) 2 (cirkev, vyznāni) <i>evaņģēliskš</i>
jams notikums	evangelijum n, (-a, pl: G -i, D -im, L -ich, I -i) <i>evaņģēlijs</i> m; e. podle sv. Lukāše sv. <i>Lūkas evaņģēlijs</i> ♦ vēřit nēčemu jako ~u <i>ticēt kaut kam kā svētajiem rakstiem</i>
eventuální eventuāls; varbūtējs;	eventualita f, (-y) <i>iespēja</i> f, <i>iespējamība</i> f, uvāžit vřechny ~y <i>apsvēřit vřas iespējas</i> ; být připraven pro vřechny ~y <i>būt gatavam jebkuram gadījumam</i>
iespējams	eventuālnē adv <i>eventuāli</i> ; (možná) <i>varbūt, iespējams</i>
evidenc/e s. 1. uzskaitē; registrā-	eventuālni adj <i>eventuāli</i> s; (možný lēž) <i>iespējami</i> s; (připadný lēž) <i>varbūtēji</i> s
cija; mít v ~i — uzskaitīt; re-	evidencje f, (-e) <i>uzskaitē</i> f, (Gpl -šu), <i>registrācija</i> f, mít (vēst) v ~i (co) <i>veikt (kā) uzskaitē, &lt;přen.&gt; paturēt prātā (ko)</i>
gistrēt; 2. filoz. evidence	evidenční adj <i>uzskaites</i> (G), <i>registrācijas</i> (G); e. číslo <i>registrācijas numurs</i> ; e. list <i>uzskaites kartīte</i>
evidenční uzskaites-; registrācijas-;	evidentně adv (očividně) <i>acimredzot</i> ; (jasně) <i>skaidri, nepārprotami</i> ;
~ lístek — uzskaites kartīte	e. (co) <i>prokázat skaidri (ko) pierādīt</i>
evoluce s. evolūcija	evidentni adj (jasný) <i>skaidri</i> s, <i>nepārprotami</i> s, <i>evidenti</i> s; (očividný) <i>acimredzami</i> s; je e., že... <i>ir skaidrs, ka...</i>
evoluční evolūcijas-; ~ teorie —	evidovaný adj <i>registrēti</i> s
evolūcijas teorija	evdlovat impf, (-uje) <i>registrēt</i> (II.), <i>uzskaitīt</i> (III.1.)
Evropān v., Eiropānka s. eiropie-	evokovat (im)pf, (-uje) (vřabavovat vřpominky) <i>atšaukt (-sauc, -sauca)*</i> (atmīņā); (nechat dojem) <i>radīt</i> (III.1.) ( <i>iespaidu</i> ); e. atmosféru <i>dētstvi atšaukt atmīņā bērnišas atmosfēru</i> ; film ~uje vřpominky na 70. lēta <i>filma atšauc atmīņā 70. gadus</i> ; obraz ~uje přestavu, že... <i>glezna rada iespaidu, ka...</i>
tis, -iete	Evropa f, (-y) <i>Eiropa</i> f, sevēni/ vřchodni/ jžiņi/ zāpadni E. <i>Ziemeleiropa f/ Austrumeiropa f/ Dienvideiropa f/ Rietumeiropa f</i> , středni E. <i>Viduseiropa f</i> , Centrāleiropa f; Rada ~y <i>Eiropas Padome</i>
evropsk/y Eiropas-; eiropieša-; eiropiešu-; ~ý rekord — Eiropas rekords; ~é jazyky — eiropiešu valodas	Evropan m, (-a, Npl -é), <b>Evropanka</b> f, (-ky, D.L -ce, Gpl -ek)
exaktní eksakts; precīzs	<i>eiropietis</i> m, (G -ša), <i>eiropiete</i> f, (Gpl -šu); ~é <i>eiropieši</i>
exaltace s. eksaltācija	evropský adj 1 (vřahující se k Evropě) <i>Eiropas</i> (G); E-ā unie <i>Eiropas Savienība</i> ; E-ā rada <i>Eiropadome</i> f; E. parlament <i>Eiropas Parlaments</i> ; E. soudni dvřr <i>Eiropas Tiesa</i> 2 (vřahující se k Evropanům) <i>eiropiešu</i> (G); ~é jazyky <i>eiropiešu valodas</i> 3 (typický pro Evropu, Evropany) <i>eiropieškš</i> ; ~é oblečeni <i>eiropiešks apģērbs</i>
exaltovaný eksaltēts	ex adv <kol.>: vypít na ex <i>izdžert vienā pagēmienā</i>
excentrický ekscentriskš	exaktní [egz-] adj <i>eksaktš</i> ; e. vědy <i>eksaktās zinātnes</i>
exekuce s. jur. 1. mantas aprakstīšana; 2. soda izpildīšana; eksekūcija	excelence f, (-e) <i>ekselenice</i> f, (Gpl -šu); Vaše/ Jeho/ Jeji E. <i>Jāsu/ Viņa/ Viņas Ekselence</i>
exekutor v. tiesu izpildītājs	exces m, (-u, L -u) <i>ekscēs</i> s m
exemplární parauga-; piemēra-; brīdinājuma-; ~ případ — parauggadījums; ~ trest — sods brīdinājuma dēļ	exekuce [egz-] f, (-e) (majetku) <i>mantas aprakstīšana</i> f; (výkon tělesného trestu) <i>soda izpilde</i> f, (Gpl -zu); (obv. hrdelnioho) <i>eksekūcija</i> f
exemplář v. eksemplārs	exekutiva [egz-, -ty/-tý-] f, (-y) (výkonnā moc) <i>izpildvarja</i> f; (výkonný orgān) <i>izpildorgān</i> s m
exil v. trimda	exemplář [egz-] m, (-e) <i>eksemplārš</i> m; ve dvou ~ich <i>divos eksemplāros</i> ; ukāzkový e. <i>paraugeksemplārš</i> m
	exhibice [egz-] f, (-e, Gpl -ti) <sport> <i>paraugdemonstrējums</i> s m
	exil [egz-] m, (-u, L -u) <i>trimda</i> f, <i>eksils</i> s m; poslat koho do ~u <i>izsūtīt ko trimdā</i> ; jít do ~u <i>doties trimdā</i> ; žít v ~u <i>dzīvot trimdā</i>

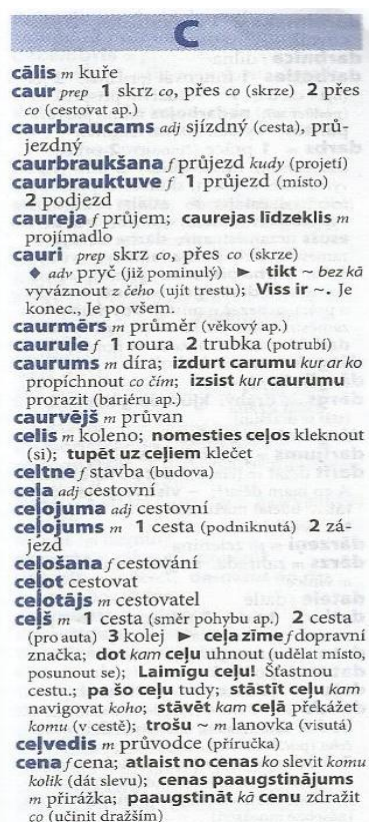
Obrázek č. 2: Ukázky zpracování hesel *etnický–exil* v ČL (1988) a ČLN (2006)

Jak patrnó z fragmentů obou slovníků, heslář doznal podstatných změn, ač obě příručky dělí necelých dvacet let. Jde tu však o více než jen přirozenou proměnu, danou bouřlivým vývojem češtiny po roce 1989; odlišnost těchto slovníků tkví spíše ve změně přístupu,<sup>17</sup> kdy je nově zohledněna zásadní vlast-

<sup>17</sup> A pochopitelně také ve snazší dostupnosti jazykových dat. Srov. však případ Česko-litevského litevsko-českého slovníku Arnoldase Piročkinase (ČLit): přestože vyšel v roce 2004, tedy pouhé dva roky před ČLN S. Nikulcevojové, zeje mezi oběma díly téměř nedozírná časová propast. Soudě podle tohoto slovníku nežijeme na počátku 21. století, ale v předinternetovém, ba předpočítačovém věku – autor zcela rezignoval na zachycení polistopadového vývoje češtiny (desítky chybějících zcela běžných a frekventovaných slov, např. *metro*, *hřeben*,

nost každého lexému, určující jeho postavení v lexikonu daného jazyka – a sice *frekvence*. Slova zastaralá, přebíraná prve zejména ze Slovníku spisovného jazyka českého, ustupují heslům z frekvenční špičky podle synchronního korpusu češtiny, objevují se kolokvialismy, přibývají a aktualizují se též kolokace.

Pro úplnost uvedme nakonec dvě konverzační příručky vydané v nedávné době: Česko-lotyšsko-litevskou konverzaci (Šelelyová et al. 2013) a česko-lotyšský díl ze série konverzací vydávaných nakladatelstvím Lingea ([Běťáková] 2012). Druhá z nich obsahuje též lotyšsko-český a česko-lotyšský slovník o celkové velikosti přibližně 6000 hesel (obr. 3), jehož kvalita je úměrná notně omezenému prostoru a nepřesahuje kvality běžných kapesních slovníků, orientovaných veskrze prakticky.



**Obrázek č. 3:** Ukázka z lotyšsko-českého slovníku vydaného v rámci česko-lotyšské konverzace (2012)

telefon, monitor, e-mail, guláš aj.) a popisuje její podobu zakonzervovanou Slovníkem spisovného jazyka českého, jenž je mu hlavním českým pramenem. – Omezme se na tomto místě na konstatování, že Piročkinasovo dílo je krokem zpátky, výslovně popírajícím veškeré moderní lexikografické trendy, a posloužilo nám mimo jiné jako příklad, jak lexikografická práce vypadat *nemá*. Podrobněji viz Škrabal 2013: 228–231.

## 2.2 LČ z hlediska typologického, jeho adresát a cíle

Obecně bývá dvoujazyčný slovník charakterizován jako „příručka mapující lexikálně-sémantické korelace a všechny další systémové vztahy mezi dvěma jazyky, s nimi pevně spjaté, která má komplexní paradigmaticko-syntagmatickou povahu“ (Čermák 1995d: 231). Lexém, ať už jakékoliv povahy, by měl být patřičně kategorizován v rámci jazykového systému (aspekt paradigmatický) a vymezen co do elementární spojitelnosti s jinými lexikálními jednotkami (aspekt syntagmatický). Třetím pomyslným pilířem dvoujazyčného slovníku je informace o ekvivalentu (aspekt sémantický),<sup>18</sup> jenž by měl optimálně usouvztažnit co nejvíce možných realizací daného lexému v textech zdrojového jazyka.

Opustíme-li nyní rovinu obecnou a definujeme-li chystaný LČ na základě běžně užívaných obsahových typologických kritérií,<sup>19</sup> pak půjde o slovník:

- *jazykový* (s velmi omezeným podílem encyklopedické informace),
- a to *dvoujazyčný* (výchozím jazykem je lotyština, cílovým čeština);
- *všeobecný* (není stratifikován geograficky či sociálně ani omezen oborově nebo zaměřen na idiolekt konkrétního autora, případně díla);
- z hlediska normativního *deskriptivní*;
- z hlediska časového *synchronní*, popisující *současný* jazyk;
- uspořádaný *sémaziologicky*, s *abecedním* řazením, bez hnízdování;
- co do média uložení se počítá s *tištěnou* podobou;<sup>20</sup>
- určen je lidskému uživateli,<sup>21</sup> má primárně *dekódovací* povahu (tj. slouží k recepci lotyšských textů českým uživatelem), počítá se proto s češtinou jako metajazykem;

---

<sup>18</sup> Srov. třídu *syntagmatics – usage – context* (Čermák 2010b: 560).

<sup>19</sup> K typologii slovníků obecně např. Čermák–Blatná 1995: 11–12; Svensén 2009: 21n.; slovníků dvoujazyčných např. Čermák 1995d: 233–234.

<sup>20</sup> O možné elektronické verzi viz dále (3.4).

- co do velikosti je to slovník *středního* rozsahu.<sup>22</sup>

Typ slovníku je do značné míry determinován cílovou skupinou uživatelů, jimž je určen. Po adresátovi svých výtvorů se logicky ptají všichni lexikografové, problematice uživatelů – jejich potřebám, jazykovým kompetencím, vyhledávacím strategiím a dalším aspektům – se věnuje zvláštní odvětví lexikografie, tzv. *user research* (Svensén 2009: 452n.). Některé jeho poznatky, byť se opírají o konkrétní jazykový materiál, lze zevšeobecnit a zohlednit i v případě LČ. Také jeho budoucí uživatel je oním „známým neznámým“<sup>23</sup> a snahy o jeho vymezení jsou spíše autorovou projekcí uživatele ideálního či univerzálního; přesto je zapotřebí pokusit se adresáta definovat, lépe řečeno je nutno nespouštět jej po celou dobu práce na slovníku ze zřetele.

Dá se očekávat dosti úzký uživatelský okruh LČ tvořený českými zájemci o lotyštinu (ať už je jejich zájem motivován jakkoliv), ještě omezenější je pak počet lotyšských zájemců o češtinu. To však nijak nesnižuje potřebnost této příručky, zvláště pro studenty jde o nepostradatelný pramen. Nezanedbatelné je ale i hledisko ryze komerční: v takovéto jazykové kombinaci nelze předpokládat stejnou míru diverzity slovníků jako v jiných případech, který by reflektovala odlišné potřeby různých typů uživatelů; naopak je nutno spojit rozdílné funkce a aspekty v rámci jedné příručky (Čermák 1995d: 232, Čermák–Blatná 1995: 12). Čtveřice, či dokonce osmice různých slovníků, implikovaná dvojí aplikovatelnou dichotomií *kódovací–dekódovací* a *pasivní–aktivní*,<sup>24</sup> je tak pouhým ideá-

<sup>21</sup> Nejde tedy o strojově čitelný slovník (*machine-readable dictionary*), sloužící k automatickému zpracování přirozeného jazyka, byť obsah databáze fungující jako podklad k LČ pro tento účel sloužit může (viz též 2.4). Srov. též Bergenholz–Tarp 1995: 236: „Dictionaries for use by humans need not contain the entire ‚lexical knowledge‘ of the language(s) in question, their purpose being to fill in linguistic competence gaps. Nor do such dictionaries need to be fully explicit, providing instead information in the form of examples which, by analogy, may serve as models. This is not so in the case of electronic dictionaries used for machine translation and other computerised natural language processing (NLP), however.“

<sup>22</sup> K vágnosti tohoto kritéria kriticky např. Čermák 1995d: 233, Svensén 2009: 25 aj.

<sup>23</sup> „*Der bekannte Unbekannte*“ (Wiegand 1977; citováno podle Svensén 2009: 452).

<sup>24</sup> V případě česko-lotyšské jazykové kombinace by se jednalo o takto funkčně vymezené slovníky:

- a) pro českého uživatele:
  1. dekódovací (pro porozumění lotyšskému textu)
  2. kódovací (pro tvorbu lotyšského textu)

lem, narážejícím nejen na překážky tržního rázu (nedostatečná poptávka po takovýchto produktech), ale venkoncem i na nedostatečně diverzifikované publikum.

Potřeby uživatelů, které dvoujazyčné slovníky plní, sumarizuje Čermák (1995d: 232) do tří hlavních okruhů, jimž odpovídají i základní typy slovníků:

1. cestování, obchod, kulturní styk (turisté, administrativní pracovníci, překladatelé a tlumočníci) – slovníky všeobecné;
2. „odborný styk vázaný na terminologii“ (překladatelé) – slovníky terminologické;
3. studium cizího jazyka (studenti, lingvisté) – slovníky studijní.

Vyhovět pokud možno všem těmto skupinám uživatelů a plnit tak tuto trojí elementární funkci si klade za cíl také LČ. Autor nezastírá ambice popsat důkladně aktuální lotyšský lexikon i s jeho specifickými složkami a jít daleko za kvalitativní standardy běžných komerčních produktů – což ovšem vyžaduje i zvýšené nároky na samotného adresáta.<sup>25</sup>

Stávající česko-lotyšské slovníky definují potřeby svých uživatelů takto:

„Slovník je zamýšlen pro ty, kteří se učí česky, čtou v tomto jazyce literaturu a tisk, překládají z češtiny do lotyštiny.“ (ČL: 5)

„Slovník je určen českým i lotyšským uživatelům, především studentům obou jazyků, překladatelům a tlumočnickům. Plní základní potřeby spojené s cestováním, obchodem

---

3. pasivní překladový (pro překlad lotyšského textu do češtiny)

4. aktivní překladový (pro překlad českého textu do lotyštiny)

b) pro lotyšského uživatele:

1. dekodovací (pro porozumění českému textu)

2. kódovací (pro tvorbu českého textu)

3. pasivní překladový (pro překlad českého textu do lotyštiny)

4. aktivní překladový (pro překlad lotyšského textu do češtiny)

Svensén 2009: 15; srov. Čermák 1995d: 233)

Reálné situaci však odpovídá jednodušší třídění, směřující první funkci s třetí a druhou funkcí se čtvrtou; ve výsledku se tak rozlišuje mezi dekodovacím (založeným na analýze textu) a kódovacím (založeným na syntéze textu) slovníkem pro Čechy a týmiž dvěma typy pro Lotyše. Přívlastky *aktivní–pasivní* bývají některými lingvisty kritizovány jako zavádějící, často se proto užívá neutrální označení *L1–L2 dictionaries* (uživatelskou mateřštinou je výchozí jazyk) a *L2–L1 dictionaries* (uživatelskou mateřštinou je jazyk cílový). (Svensén 2009: 16)

<sup>25</sup> Že je kompetence uživatelů při práci se slovníkem neméně důležitá než schopnost lexikografa vyhovět jejich potřebám, je nábídní. „Users may depend on the dictionary for giving them the contextualised information they need, but in the end it is the user who makes the decision and applies the information. In other words, dictionary use needs to be seen as an interactive occupation in which both participants, the dictionary and its user, need to be highly qualified and know their role in the decision-making process.“ (Varantola 2003: 231)

i s dalšími praktickými nebo kulturními styky a pomůže každému, kdo se zajímá o jazyk druhého národa.“ (ČLN: zadní strana obálky)

Těmto východiskům odpovídá i samotný způsob zpracování hesel: ČL je orientován primárně na lotyšského uživatele a je pasivní povahy, zatímco ČLN je koncipován jako obousměrný, snažící se posloužit oběma skupinách mluvčích, nicméně má prvotně charakter kódovací (Nikuļceva 2013: 214).<sup>26</sup>

V tomto ohledu je LČ kompromisem mezi oběma přístupy: na obecné rovině si jeho adresáta definujeme jako vzdělaného, poučeného uživatele, do jisté míry znalého obou jazyků, jemuž je předkládáno rozšířené jádro lotyšského lexikonu. Ze slovníku bude těžit především český uživatel, jenž by měl s jeho pomocí dekodovat lotyšské texty. Příručku ovšemže využijí i lotyšští uživatelé při produkci českých textů; ti ale musí být eventuálně připraveni (zvláště při nižší úrovni znalosti češtiny) konzultovat získané informace s jinými zdroji. LČ však není primárně koncipován jako slovník obousměrný, už pro mnohem vyšší nároky na prostor,<sup>27</sup> ale též na zpracování (skepticky vůči bifunkčním slovníkům též Čermák 1995d: 246).

### 2.3 Lexikografické zdroje pro tvorbu LČ

Logickým východiskem pro korpusově orientovanou lexikografii jsou po materiálové stránce jazykové korpusy. Aktuální situaci jsme nastínili už v předminulé kapitole, překonat onu „závažnou překážku“ je třeba kompenzovat jinými zdroji: materiálová základna se tak přirozeně rozrůstá o další prameny, využívané tradiční lexikografií.

---

<sup>26</sup> Potvrzeno samotnou autorkou (Nikuļceva, osobní komunikace).

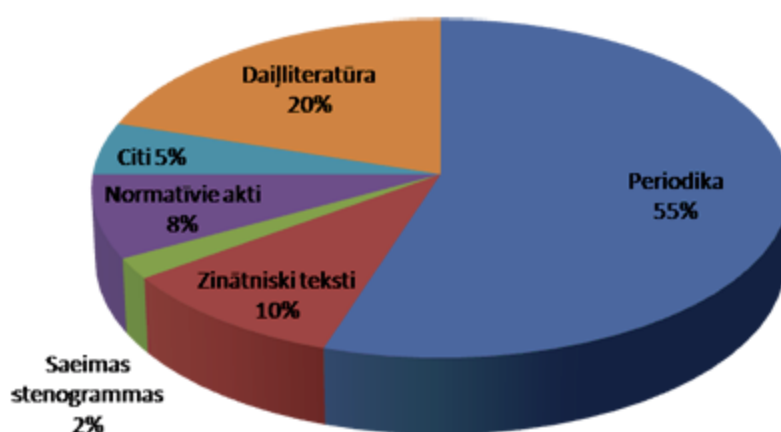
<sup>27</sup> „[...] genuine nondirectional dictionary necessitates considerable repetition of information, which greatly increases its volume. Even more inconveniently for the user, the entries – especially those for frequent words, which tend to be long anyway – may look intimidating, the extra information in the ‘other’ language slowing down the consultation process, and possibly culminating in abandoning the search together.“ (Adamska-Sałaciak 2006: 92).



### 2.3.1 Korpusy

2.3.1.1 V Lotyšsku probíhá budování jazykových korpusů, v čele s reprezentativním synchronním, od roku 2005 na půdě Laboratoře umělé inteligence Institutu matematiky a informatiky Lotyšské univerzity v Rize.<sup>28</sup> Po pilotním korpusu *ledus*<sup>29</sup> a prvotní verzi vyváženého korpusu psané lotyštiny *miljons*<sup>30</sup> byl v roce 2013 zveřejněn korpus lotyštiny LVK2013 (dále LVK), obsahující 5 588 489 pozic (z toho 1 017 428 tvoří interpunkce), 296 518 slovních tvarů a 116 605 lemmat, morfologicky označovaný a lemmatizovaný. Žánrově je složen takto: 55 % textů představuje publicistika, 20 % beletrie, 10 % odborná literatura, 8 % právní předpisy, 2 % stenografické záznamy jednání poslanecké sněmovny a zbývajících 5 % jiné texty.<sup>31</sup>

Žánrové složení vystihuje následující tabulka:



Obrázek č. 4: Žánrové složení korpusu *Latviešu valodas korpuss 2013* (převzato z <http://korpuss.lv>)

<sup>28</sup> *Latvijas Universitātes Matemātikas un informātikas institūta Mākslīgā intelekta laboratorija – Allab* (<http://ailab.lv/>). Toto pracoviště však shromažďuje texty za účelem jejich strojového zpracování již od konce 80. let minulého století (Skadiņa – Vasiljevs 2013: 456).

<sup>29</sup> Text první kapitoly z románu Paulse Bankovskise *Plāns ledus* (Tenký led) z r. 1998. Jde o ručně značkový a lemmatizovaný, trénovací korpus s 16 746 pozicemi (bez interpunkce necelých 14 tisíc pozic, 5157 slovních tvarů, 3370 lemmat). Blíže viz bakalářská práce K. Levāneové (Levāne 2001).

<sup>30</sup> Verze 1.0 s 1 109 968 pozicemi (bez interpunkce cca 1 milion, 115 730 slovních tvarů) je netagovaná, stejně jako 2.0, navzdory stejnému názvu čtyřikrát větší (4 409 097 pozic, bez interpunkce cca 3,5 milionu, 251 901 slovních tvarů). Verze 2.0m, ač menší (3 885 303 pozic, bez interpunkce 3 194 856 pozic, 237 980 slovních tvarů, 62 809 lemmat), už je morfologicky označovaná.

<sup>31</sup> Podrobnější statistika je dostupná online na <http://korpuss.lv/uzzinass/statistika.html>. Kritéria výběru textů popisuje Levāne-Petrova 2012.

Tento korpus je pro nás základním zdrojem, kromě něj pracujeme rovněž s korpusem *Latviešu valodas tīmekļa korpus* (dále LVTK), zkompilevaným z lotyšských webových stránek, o velikosti 122 628 720 pozic (bez interpunkce cca 97 milionů). Ostatní korpusy<sup>32</sup> využíváme pro jejich přílišnou žánrovou, diachronní či diatopickou specifickou jen sporadicky a v omezené míře. Většina lotyšských korpusů je zpřístupněna pomocí korpusového manažeru Manatee s uživatelským rozhraním Bonito, některé jsou dostupné online přímo ze stránek <http://korpus.lv/>.

Užívání lotyšských korpusů však má svá úskalí, která negativně ovlivňují především analytickou fázi lexikografické práce. Hlavním záporem je malá velikost LVK: na takto omezeném objemu dat lze zkoumat jevy pouze centrální, periferní úkazy se často vyskytují v několika málo výskytech (včetně nízké variantnosti flektivních forem daného lexému), případně se neobjevují vůbec. LVK by tak mohl posloužit jen jako materiálová základna pro mnohem menší překladový slovník (slovník výkladový lze pak vyloučit úplně; srov. Zuicena–Migla 2014), než je chystaný LČ. LVTK svým rozsahem sice odpovídá dílčímu českému korpusu řady SYN (na podkladu korpusu SYN2000 byl zhotoven ČLN o rozsahu cca 23 tisíc hesel), avšak je nelemmatizovaný: je v něm tudíž možno vyhledávat pouze podle slovních tvarů či tagů, přičemž značný podíl textů bez

---

32

- korpusy *saeima-1.0* a *saeima-2.0*, tvořené záznamy z jednání lotyšské poslanecké sněmovny (*Saeima*) z let 1993–2008; sloužily jako materiálová základna do LVK (viz výše uvedený 2% podíl). Celková velikost přes 27 milionů pozic; netagováno, nelemmatizováno;
- korpus latgalských textů *MuLa-1.0* (*Mūsdienu latgaliešu tekstu korpus*) o celkové velikosti přes 1 milion pozic; netagováno, nelemmatizováno;
- paralelní litevsko-lotyšský korpus *LiLa* (*Lietuviešu-latviešu-lietuviešu paralēlo tekstu korpus*) o celkové velikosti přes 1 milion pozic; netagováno, nelemmatizováno;
- neprohledávatelný korpus faksimilií starých lotyšských textů (*Latviešu valodas seno tekstu korpus*), tvořený lotyšskými písemnými zdroji z 16., 17. a počátku 18. století o celkové velikosti přes 900 tisíc pozic (bližší informace na <http://www.korpus.lv/senie/about.htm>).
- řada korpusů *emuari-0*, *emuari-40*, *emuari-80* a *emuari-90*, zachycující aktuální lotyšskou blogosféru o celkové velikosti cca 12 milionů pozic, tagováno, lemmatizováno. De facto však jde o čtyři subkorpusy jednoho korpusu, vyčleněné na základě kvalitativních měřítek, číslce v názvech odkazují k minimální hodnotě poměru rozpoznávaných slov. Bližší informace viz bakalářská práce M. Laizāne (Laizāns 2015).

Mimoto jsou na stránkách <http://korpus.lv/> odkazy, bohužel nefunkční, na syntakticky značkováný korpus (treebank) lotyštiny a korpus gramatických a sémantických valencí nejčastějších lotyšských sloves. Za zmínku stojí rovněž neexistence jakéhokoliv mluveného korpusu (nepovažujeme-li za něj korpusy *saeima*), k jeho přípravám viz Auziņa 2010.



diakritiky hledání ještě komplikuje. Kvalitu korpusu dále snižují, a to v některých případech až na hranici únosnosti, hojné cizojazyčné fragmenty, šablony (tzv. *boilerplate*) a neméně častá defektní místa (rozdělená slova či naopak slitky, špatné formátování aj.), i navzdory proklamovanému odstranění duplicit se v korpusu příliš často objevují totožné či minimálně se lišící věty, což výrazně zkresluje frekvenční statistiku.<sup>33</sup>

Druhou komplikací je nejednotné morfologické značkování v různých lotyšských korpusech,<sup>34</sup> vedoucí mimo jiné k odlišné anotaci týchž jednotek a sporné lemmatizaci. Tak je ve dvou námi využívaných korpusech odlišně chápáno např. sousloví *droši vien* „asi, nejspíš“: zatímco v LVTK se v důsledku nepřítomné lemmatizace objevuje v konkordancích obou složek, a to celkem 3477krát, v LVK se po zadání tvaru *droši* ani *vien* nezobrazí (opačná složka tak není ani součástí kolokačního profilu daného tvaru), neboť je lemmatizováno jako víceslovná jednotka a k jeho vyhledání je potřeba zadat dotaz `[lemma="droši vien"]`, na nějž získáme 464 výskytů. To však nemusí být uživateli zřejmé, nejasná jsou též kritéria, na základě nichž byly víceslovné jednotky definovány.<sup>35</sup>

Morfologická anotace lotyštiny se opírá o „gramatický analyzátor“ SemTi-Kamols<sup>36</sup> a je plně automatická, ačkoliv se počítá též s možností manuálního doplnění jednotlivých tagů a vydělení specifických tříd lexémů.<sup>37</sup> Lotyšský tagset, jejž nástroj SemTi-Kamols využívá, je odvozen od výsledků

---

<sup>33</sup> Sběr obsahu webových stránek probíhal automaticky pomocí dnes již nefunkčního internetového vyhledávače <http://search.latnet.lv/>, taktéž automaticky došlo k eliminaci duplicitních vět. Robustnější filtrovací nástroje by výslednou kvalitu korpusu byly výrazně zlepšily (srov. Suchomel – Pomikálek 2012).

<sup>34</sup> <http://korpuss.lv/uzzinas/instrukcija.html>

<sup>35</sup> Dotaz `[lemma=". * ."]` nalezne v LVK 19 337 výskytů, z nějž lze vygenerovat seznam celkem 855 víceslovných lemmat, po nezbytném odfiltrování numerických lemmat (nejčastější 10 000 se 72 výskyty) a slov rozdělených či proložených (např. *hapax groteska*) pak pouhých 49 lexémů. Většina z nich bude zařazena do hesláře LČ jako víceslovná lemmata (viz dále 3.1.2.4.1). Jen stěží však může jít o více než frekvenční špičku víceslovných jednotek v lotyštině.

<sup>36</sup> <http://www.semti-kamols.lv/>

<sup>37</sup> Týká se to výhradně substantiv (hodnota *n* na 1. pozici tagsetu), konkrétně pozic č. 3 (rod, hodnota *k* pro tzv. „kontextuální rod“, tj. obourodá substantiva), 4 (číslo, k odlišení singularií tantum od obvyklých singularových tvarů, resp. plurálií tantum od plurálových), 5 (pád, k rozlišení ustrnulých genitivních forem a běžných genitivů) a zčásti též 2 (druh substantiva, částečně jsou automaticky rozlišována propria od apelativ). Tato možnost však zůstává dosud nevyužita, a vyčlenit tyto zvláštní skupiny slov proto nelze.

mezinárodního projektu MULTEXT-East (z něhož vzešel i český tagset používaný u korpusů řady SYN<sup>38</sup>) a jeho zevrubný popis lze nalézt na [http://www.semti-kamols.lv/doc\\_upl/TagSet.html](http://www.semti-kamols.lv/doc_upl/TagSet.html) (srov. též Levāne – Spektors 2000, Levāne 2001). Navzdory 93,6% úspěšnosti taggeru uváděné samotnými tvůrci (Paikens – Rituma – Pretkalniņa 2013) dochází k případům, kdy se lexikograf na morfologickou anotaci v lotyšských korpusech nemůže spolehnout (za všechny viz pozn. 196).

Aktuálně největší dostupný webový korpus lotyštiny – *lvTenTen* (o celkové velikosti 667 668 379 pozic) – však vznikl mimo rámec AILab, na brněnské Masarykově univerzitě: v roce 2014 byl postahován crawlovacím nástrojem SpiderLing (Suchomel – Pomikálek 2014) jakožto člen korpusové rodiny *TenTen* (Jakubíček et al. 2013). Prvotní verze byla převedena do formátu UTF-8, vyčištěna a zbavena duplicit, korpus zůstal prozatím neoznačkován, zato nabízí možnost slovních profilů (funkce Word Sketches<sup>39</sup>), kterou jiné korpusy neposkytují. V tom spatřujeme jeho největší přednost, byť je naše zkušenost s tímto korpusem malá, omezená na pouhých pár sond.

2.3.1.2 Dalším korpusovým zdrojem je česko-lotyšská část paralelního korpusu InterCorp (dále IC), která v současné verzi (8.0 z června 2015) obsahuje přes 32 milionů pozic: původní beletristické, ručně zarovnané jádro (aktuálně 1 336 000 pozic) bylo, podobně jako u ostatních jazyků v IC, rozšířeno o kolekce automaticky zpracovaných textů: v případě lotyštiny konkrétně o právní texty EU z korpusu *Acquis Communautaire* (18 744 000 pozic), zápisy jednání Evropského parlamentu z let 2007–2011 z korpusu *Europarl* (11 688 000 pozic) a balíček filmových titulků z databáze *OpenSubtitles* (280 000 pozic).<sup>40</sup>

---

<sup>38</sup> Erjavec 2004, viz též <http://nl.ijs.si/ME/Vault/V3/msd/msd.pdf>.

<sup>39</sup> Výhledově se dá jistě počítat i s dvoujazyčnými slovními profily včetně česko-lotyšské kombinace (Kilgariff 2013: 92n.).

<sup>40</sup> Srov. stav v roce 2011, ver. 3.1 (Škrabal 2011: 263). Obecně k IC Čermák–Rosen 2012.

Ač značně nereprezentativní, je velmi hodnotným zdrojem, neboť jako jediný bezprostředně nabízí české ekvivalenty lotyšských lexémů. Jirásek (2011: 53) upozorňuje na riziko, že u korpusů podobného objemu a složení mohou převážít idiolektické prvky nad běžnou jazykovou praxí: „Abychom korpusem nabízený ekvivalent mohli považovat za dostatečně spolehlivý, musí se v něm vyskytovat opakovaně, v různých textech a u různých autorů. [...] u slov s jedním nebo dvěma výskyty v korpusu je výskyt jejich ekvivalentu prakticky věcí náhody, nemůžeme tu uplatnit statistické hledisko.“ (ibid.) Zohlednit je třeba i možnost vyhnout se přímé ekvivalentaci daného lexému, již překladatelé hojně využívají, a připustit musíme i překladatelská řešení sporná či vyloženě chybná: ta ve slovníku ovšemže nemají své místo.

Hlavními slabinami lotyšské části IC je tedy jeho velikost<sup>41</sup> a skladba textů – přesněji řečeno minimální podíl lotyšských originálů,<sup>42</sup> které by pro naše účely byly ideální – a rovněž dosud neimplementovaná morfologická anotace, což výrazně omezuje možnosti prohledávání.

Značně limitované objemem a složením dat jsou nicméně jiné volně přístupné česko-lotyšské paralelní korpusy.<sup>43</sup>

2.3.1.3 Z lexikografické perspektivy přínosná je i nová aplikace Treq,<sup>44</sup> využívající metody automatického zarovnávání paralelních textů na úrovni slov (stochastický program GIZA++; Och–Ney 2003) k vyhledávání potenciálních lexikálních překladových ekvivalentů. Po zadání příslušného lexému (v nelem-

---

<sup>41</sup> Srov. zkušenost z chorvatsko-české části IC: „Ukázalo se, že pokud nechceme zůstat na úrovni kapesních slovníků, potřebujeme paralelní korpus o velikosti alespoň kolem 10 miliónů textových slov. Takovýto rozsah korpusu nám umožní vcelku spolehlivě zpracovat ekvivalenty pro slovník středního rozsahu. Pro slovník většího rozsahu však může sloužit jen jako orientační pomůcka, nikoli hlavní zdroj ekvivalentů.“ (Jirásek 2011: 55) Slovníkem středního rozsahu je tu přitom myšlen slovník obsahující cca 20 tisíc hesel, uvádějící pouze „významy typické a převládající v každodenní komunikaci“, slovník většího rozsahu by měl pojmut cca 50 tisíc hesel (ibid.: 45).

<sup>42</sup> V jádru jde o 5 románů, 1 knihu pohádek a 1 kratší esejistický text, v kolekci *Europarl* o 268 opusů od 16 různých původců. Celková velikost takovéhoho subkorpusu činí 389 489 textových pozic, tj. cca 0,9 % z celkového objemu dat v lotyšské části IC (v samotném jádru pak poměr lotyšských originálů tvoří cca 20,2 %).

<sup>43</sup> Např. otevřený paralelní korpus *Opus* (<http://opus.lingfil.uu.se/>) či nástroj *Linguee* ([www.linguee.com](http://www.linguee.com)), spojující redakční slovník s vyhledávačem.

<sup>44</sup> <http://treq.korpus.cz/>. Spuštěna byla spolu s verzí IC 8.0.

matizované lotyšské verzi konkrétního tvaru) získáme soupis kandidátů na české protějšky i s jejich frekvencí (absolutní i vyjádřenou procentuálně), která je přímo úměrná pravděpodobnosti, že daný kandidát je funkčním ekvivalentem. Výhodou je možnost kterýkoliv z nich rozkliknout a ověřit si jeho výskyt v kontextu, snadněji a rychleji se tak vytřídí ekvivalenty relevantní od těch zavádějících. U velkých a žánrově pestrých korpusů má zajisté smysl uvádět frekvenci ekvivalentních párů zvlášť pro různé typy textů (beletrie, publicistika, právnické texty, filmové titulky) (Rosen – Adamová – Vavřín 2014: 178), pro lotyšský korpus je toto členění distinktivní zatím jen u nejfrekventovanějších lexémů.

Stejný princip, byť za pomoci jiných nástrojů, byl už v českém prostředí využit např. v případě chorvatštiny (databázová aplikace Microsoft Access; Jirásek 2011) a litevštiny (nástroj NATools,<sup>45</sup> používaný původně pro extrakci terminologických slovníků; Skoumalová 2010). Oba autoři se shodují na tom, že slovníky takto automaticky extrahované z paralelního korpusu jsou pouhým východiskem pro následnou slovníkářskou práci, mohou ji však lexikografovi nemálo ulehčit. Další zlepšení výsledků lze očekávat spolu se zvětšením objemu dat a žánrové pestrosti použitých textů a také se zpracováním víceslovných výrazů, jež by zohlednilo případy asymetrické ekvivalentace, tedy s rozšířením metody párující lexémy s jejich ekvivalenty striktně v poměru 1:1.

2.3.1.4 V neposlední řadě je z korpusových zdrojů třeba zmínit též synchronní korpusy češtiny, především aktuální referenční SYN2015 (přibližně stomilionový) a nereferenční SYN o velikosti cca 2,232 miliardy pozic. Ty jsou pro autora slovníku důležité pro ověřování životnosti a přirozenosti českých ekvivalentů, ať už nabízených v paralelních korpusech nebo jím samotným vy-

---

<sup>45</sup> Autor by zde rád poděkoval dr. Haně Skoumalové za pomoc při aplikaci tohoto nástroje na lotyšsko-českou část IC.

vozených, v případě vícerych ekvivalentů pak k jejich správnému řazení v rámci hesla.<sup>46</sup>

### 2.3.2 Slovníky

Tradičním lexikografickým pramenem zůstávají slovníky, především výkladové a jiné překladové slovníky s lotyštinou v roli zdrojového jazyka. V našem případě, kdy je objem korpusových dat omezený, se stávají nepostradatelným zdrojem, je však nutno mít na zřeteli nejméně trojí riziko. Zaprvé je popis ve slovnících vždy více či méně zastaralý, neodpovídající aktuálnímu úzu, a v mnohém i chybný. Dále existuje nebezpečí, že autor bude cizím zdrojem ovlivněn a bude mít (třeba jen bezděčnou) tendenci přizpůsobovat zpracování lexému popisu ve zdrojovém či jiném cílovém jazyce než čeština, přičemž každý popis je založen na nějakém konkrétním teoretickém pozadí (Sinclair 1991: 37n.), jež nemusí odpovídat námi zvolenému východisku.<sup>47</sup> Konečně poměrně častá je praxe přejímání lexikografických řešení ze starších zdrojů bez dalšího ověřování v nezávislých pramenech; ani přítomnost lexému v hesláři vícero slovníků tak není zárukou, že dané slovo je součástí živého jazykového úzu (Svensén 2009: 60 mluví o tzv. *ghost words*, tento koncept však lze rozšířit též na kolokace i celé příkladové věty, které, ač notně „šustí papírem“, objevují se znovu a znovu).<sup>48</sup>

V dnešní době významně roste podíl elektronických slovníků, mnohé tištěné příručky jsou elektronizovány, jiné mají podobu pouze virtuální. Primárním zdrojem je pro nás webová stránka <http://tezaurs.lv/>, na níž je soustře-

---

<sup>46</sup> Jde o velmi cennou, byť dostatečně nedocněnou devizu korpusů cílového jazyka. Lexikograf by se ani v této fázi neměl spoléhat na svou intuici, ale hledat objektivní data v korpusech. Bezpečně rozsouditelné jsou snad jen případy, kdy je některý z konkurentů příznakový (*který × jenž*), oproti tomu např. frekvenční poměr synonym *stínítko* a *stínidlo* [lampy] (protějšky lotyšského lexému *abažūrs*) lze jen sotva vyvodit intuitivně.

Srov. též možné využití aplikace Morfio: <http://morfio.korpus.cz/QhczofDB>.

<sup>47</sup> Pochopitelně také každá anotace korpusových dat nutně vychází z konkrétní teorie a je tedy svým způsobem jejich interpretací. Lexikografovi by tedy měla zůstat pouhým vodítkem a nástrojem, nikoliv nedotknutelným axiomem.

<sup>48</sup> Kopeckij 1974: 32–33: „U velkých slovníků tradice slovníkových předchůdců někdy pomáhala na svět jejich lexikálnímu balastu, vyumělkovaným konstrukcím a nikde nedosvědčeným slovům nebo výrazům, které byly přejímány ze slovníku do slovníku.“

děň obsah bezmála tří stovek (sic!) lotyšských lexikografických a encyklopedických děl.<sup>49</sup> Využíváme především tyto slovníky:

2.3.2.1 Ze slovníků výkladových je pro nás páteřním zdrojem Slovník současného lotyšského jazyka (*Mūsdienu latviešu valodas vārdnīca*, MLVV), nejnovější výkladový slovník lotyštiny o velikosti téměř 45 tisíc hesel, vznikající v letech 2001–2014<sup>50</sup> a dostupný online.<sup>51</sup> Čerpá z lexikální kartotéky, zčásti převzaté po starších lexikografických projektech, z internetových a také korpusových dat a z jiných slovníků a encyklopedií, které v Lotyšsku vyšly v posledních letech, východiskem však jsou dřívější výkladové slovníky. Oproti nim lze u MLVV vedle snahy o ideologicky neutrální definice některých slov zaznamenat četnější informace encyklopedického rázu (latinské názvy botanických a zoologických termínů, etymologie slov cizího původu aj.) a rozšíření hesláře o některá propria či zkratky. Mimoto jsou vybraná hesla komentována z hlediska jazykové správnosti a kultury jazyka (slovník tak částečně plní i roli ortografické a ortoepické příručky). Nesoulad s kodifikovanou normou však předem nevylučuje daný lexém z hesláře; MLVV se snaží o zachycení celkového obrazu soudobé lotyštiny v různých jejích vrstvách<sup>52</sup> a plně tak konvenuje našim představám o charakteru LČ (nehledě na podobný počet hesel). Pomíne-li drobné nepřesnosti v jednotlivých heslech, včetně nesouladu s popisem v pokynech pro uživatele, a dílčí nekonzistence napříč heslářem,<sup>53</sup> je největší

---

<sup>49</sup> Aktuální verze (zima 2016), zahrnující 282 zdrojů (klíčovým je LLVV s více než 25% podílem), obsahuje 256 423 hesel a 105 499 synonym [údaj aktuální k 30. 3. 2016]. Kompletní soupis zdrojů viz

<http://tezaurs.lv/#/avoti>.

<sup>50</sup> Přípravné práce však započaly již v lednu 1997, bezprostředně po vydání posledního svazku LLVV (Zuicena 1998: 122, Zuicena – Migla 2014).

<sup>51</sup> <http://www.tezaurs.lv/mlvv/>

<sup>52</sup> Příznačný je již nový titul slovníku, obsahující neutrální přívlastek *mūsdienu* ‚současný‘, jímž nahradil dřívější *literārās* ‚spisovný‘ (LLVV). Bohatší spektrum tvoří též ilustrativní příklady užití lexémů: nejsou už jako dříve čerpány téměř výhradně z beletrie (proti Ertlovu konceptu „dobrého autora“ je Čermák 1995c: 107), ale zejména z aktuální publicistiky a z internetových zdrojů a korpusových dat. Viz též Zuicena 1998: 124 a Zuicena 2012: 236.

<sup>53</sup> Jejich souhrn podává např. ve své diplomové práci Jērane 2013. Vzhledem k elektronické povaze MLVV se dají všechny výtky nicméně snadno zapracovat, s fází oprav a korektur se počítá pro roky 2014–2017, dále pak s pravidelnou revizí slovníku jednou za čtyři roky. Stejně tak je plánováno vylepšení slovníku po technické stránce a jeho rozšíření o další materiál, včetně audiovizuálního (Zuicena–Migla 2014).

slabinou slovníku jeho minimalistické uživatelské rozhraní: absence samotného hesláře či jakéhokoliv našeptávače, ba i klávesnice s lotyšskou diakritikou velmi stěžuje vyhledávání hesel.

Dvojice starších výkladových slovníků – jednosvazkový Slovník lotyšského jazyka (*Latviešu valodas vārdnīca*, LVV), respektive jeho revidované a doplněné vydání z r. 2006 o velikosti cca 30 tisíc hesel,<sup>54</sup> a osmisvazkový Slovník spisovného lotyšského jazyka (*Literārās latviešu valodas vārdnīca*, LLVV), vydávaný v letech 1972–1996 a čítající přes 64 tisíc hesel – byla vytěžována zvláště v prvotní fázi naší práce, než byl dokončen projekt MLVV; dnes nám slouží pouze jako doplňkový materiál. Ke starším slovníkům včetně kanonického díla lotyšské lexikografie, čtyřsvazkového Slovníku lotyšského jazyka<sup>55</sup> Kārlise Mīlenbahse a Jānise Endzelīnse z let 1923–1946 (MEV, pro jejich neaktuálnost nepřihlížíme.

2.3.2.2 Mezi překladovými slovníky se našim představám o LČ nejvíce blíží lotyšsko-polský slovník (LP) autorského kolektivu pod vedením Albertse Sarkanise, vydaný v roce 2011 a obsahující okolo 48 tisíc hesel (největší v této jazykové kombinaci). V lecčems může sloužit za vzor: z řady jiných překladových slovníků vystupuje moderním lexikonem včetně výrazů do hesláře běžně nezahrnovaných (slangismy, vulgarismy aj.), uživatelsky vstřícným přístupem a některými netradičními prvky (např. netradiční použití druhé barvy pro polské ekvivalenty). Nelze však přehlížet ani jeho neduhy: jednak se tvůrci neopírali o korpusová data, berouce mnohdy zavděk popisem ve starších slovnících (to- tožné kolokace, přebírané omyly), jednak jejich velký počet (celkem 14) nepřispívá ke konzistenci zpracování – editor tu byl postaven před nelehký úkol sjednotit příliš mnoho individuálních lexikografických pojetí a řešení, což se ne

---

<sup>54</sup> První vydání slovníku pochází z roku 1987, reedice z roku 1998, obě verze obsahují kolem 25 tisíc hesel. Heslář byl rozšířen o nové lexémy, definice některých hesel byly podobně jako v případě MLVV zbaveny dobového ideologického témbu.

<sup>55</sup> Dostupný online na <http://www.tezaurs.lv/mev/>.

vždy podařilo. I tak je možno hodnotit výslednou podobu slovníku pozitivně a využívat jej, už pro blízkost češtiny a polštiny, jako jeden z důležitých zdrojů.

Inspirativní mohou být také další překladové slovníky, zvláště ty jasné dekódovacího charakteru, jež jsou však v menšině, zvláště v kombinacích s velkými jazyky; četné slovníky lotyšsko-anglické,<sup>56</sup> lotyšsko-ruské<sup>57</sup> či lotyšsko-německé lze využívat jen v omezené míře. Slovník lotyšsko-litevský (LL) je pro nás nevyhovující praxí hnízdování, na několika místech v této práci k němu však přihlížíme. – Přehled o celkové produkci překladových slovníků v Lotyšsku v období 1991–2010 včetně vyčerpávající bibliografie podává Baldunčiks 2012, zčásti též Jansone–Veisbergs 2013 či Schmidt 1991).

Do určité míry lze české protějšky lotyšských lexémů ověřovat ve dvojici česko-lotyšských slovníků, které byly stručně představeny již v kap. 2.1. Z elektronické podoby rukopisu ČLN byla pro naše účely počítačově vytvořena obrácená verze slovníku:<sup>58</sup> každý první lotyšský ekvivalent u všech číslovaných významů byl změněn v lemma, doplněné o zbytek původního hesla. Získali jsme tak nejen heslář o více než 31 tisících položkách, ale rovněž jeho české protějšky spolu s kolokacemi.

**absenc|e** f, (-e) **1** (nepřítomnost) *prombūtn|e* f; zastupovat koho v době jeho a. *aizvietot ko viņa prombūtnes laikā* **2** (nedostavení se) *neierašān|ās* f, *kavējum|s* m; neomluvená a. *neattaisnots kavējums* **3** (jevu, vlastnosti) *trūkum|s* m; a. smyslu pro humor *humora izjūtas trūkums*

**prombūtn|e** f; zastupovat koho v době jeho a. *aizvietot ko viņa prombūtnes laikā* ==> **absenc|e** f, (-e) - (nepřítomnost)

**neierašān|ās** f, **kavējum|s** m; neomluvená a. *neattaisnots kavējums* ==> **absenc|e** f, (-e) - (nedostavení se)

**trūkum|s** m; a. smyslu pro humor *humora izjūtas trūkums* ==> **absenc|e** f, (-e) - (jevu, vlastnosti)

Objevil-li se lotyšský lexém jako ekvivalent více českých lemmat, vygenerováno bylo i vícero hesel se stejným lemmatem na lotyšské straně.

**trūkum|s** m; a. smyslu pro humor *humora izjūtas trūkums* ==> **absenc|e** f, (-e) - (jevu, vlastnosti)

<sup>56</sup> V čele s Novým lotyšsko-anglickým slovníkem (*Jaunā latviešu-angļu vārdnīca*, Veisbergs 2005) s cca 62 tisíci hesel, autorem soustavně doplňovaného.

<sup>57</sup> Dvojsvazkový LR z r. 1981 (reedice v r. 1996) s cca 53 tisíci hesly, z nichž mnohá jsou poplatná dobové ideologii, nám může pomoci podchytit základní lexémy příznačné pro sovětskou dobu (v LČ důsledně označované markerem <sov>, viz 3.2.6). Z novějších slovníků okrajově využíváme ilustrovaný LKIV (cca 16 tisíc hesel).

<sup>58</sup> Autor práce by zde rád poděkoval ing. Zdeňku Velískovi za zhotovení speciálního programu, který umožnil její vznik.



**trūkum|s m; pro n. času laika trūkuma dēļ; n. dūkazū pierādījumu trūkums; nemá v ničem n. viņam nekā netrūkst** ==> nedostat|ek m, (-ku, Lpl -cīh) -

**trūkum|s m, vain|a f, defekt|s m; odstranit ~ky novērst trūkumus** ==> nedostat|ek m, (-ku, Lpl -cīh) - obv. pl ~ky (vada)

**trūkum|s m, nabadzīb|a f; Žīt v ~ku dzīvot trūkumā** ==> nedostat|ek m, (-ku, Lpl -cīh) - (nouze)

**trūkum|s m; (nesnáz) grūtīb|as f, pl; časová/ bytovā t. laika/ dzīvokļu trūkums; finanční t. finansiālas grūtības** ==> tīs|eñ f, (-nē) - (nouze)

**trūkum|s m, nabadzīb|a f; Žīt v ~i dzīvot trūkumā (nabadzībā); mīt, trpēt nouzī ~i ciest trūkumu** ==> nouz|e f, (-e) - (bīda) β *třít bīdu s ~i būt galīgā postā*

**trūkum|s m, nabadzīb|a f; Žīt v ~i dzīvot trūkumā (nabadzībā); mīt, trpēt nouzī ~i ciest trūkumu** ==> nouz|e f, (-e) - (bīda) β *třít bīdu s ~i būt galīgā postā*

**trūkum|s m; bīda o peníze naudas trūkums** ==> bīd|a f, (-y) - (nedostatek)

Takto solidní báze poskytuje nejen výčet českých ekvivalentů, ale též představu o jejich sémantickém rozvrstvení; tímto způsobem získaná data by nicméně měla být konfrontována s korpusovými daty a popisem v lotyšských výkladových slovnících, což se týká především kolokací a frazémů.<sup>59</sup>

2.3.2.3 Dále čerpáme z nejrozličnějších speciálních slovníků: frazeologického (LFV), synonymického (LVSiV), ortoepického a ortografického (LVPPV), terminologických (všeobecné: např. BLJuT, i specializované: např. VPSV aj.; k pokrytí základní oborové terminologie dobře slouží též středoškolské učebnice a přehledy), slovníků cizích slov (ISV, Rozenvalde 2004, Baldunčiks 2007), eponym (Bankava 2007), křestních jmen (Siliņš 1990), zeměpisných jmen (LPA), slangu a argotu (LVSV, LVŽV, Kavalieris 2002), zkratk (Bankavs 2003) a mnoha jiných příruček (např. Ruberte 2004, Kursīte 2007, Kursīte 2009, Šustere 2009 aj.). K ověřování frekvenční špičky lze využít frekvenčního slov-

<sup>59</sup> K možnostem automatického převrácení dvoujazyčných slovníků se vyjadřuje spíše skepticky Svensén (2009: 62), v nejlepším případě jde podle něj o podkladový materiál, „jenž musí následně projít výběrem, doplněním, novou redakcí atd.“; k zanedbání této fáze očividně došlo u automaticky převráceného lotyšsko-estonského slovníku (LE), viz recenze Ojārs Buše (Bušs 2015). Ideální je tento nástroj v případě terminologických slovníků, obecněji řečeno u monosémantických lexémů, jež tvoří ve všeobecně zaměřených překladových slovnících cca 35–45 % hesláře (Corda et al. 1998: 442). Srov. též zkušenost s automatickým obrácením anglicko-lotyšského slovníku (Veisbergs 2003: 169–174).

níku lotyštiny (LVBV) a přehledy lexikálního minima (Grünberga et al. 1997, Kuzina 1998a).<sup>60</sup>

2.3.2.4 Z českých slovníků vycházíme především ze Slovníku spisovné češtiny (SSČ) a Nového akademického slovníku cizích slov (NASCS), pro ověřování synonymických ekvivalentů pak ze Slovníku českých synonym (Palašianský 2000) a Tezauru jazyka českého (Klégr 2007).

### 2.3.3 Další lingvistické zdroje

Lexikon jazyka není samozřejmě popsán pouze v lexikografických a encyklopedických příručkách, ale zčásti také v gramatikách, lingvistických monografiích a studiích, ba i v populárně naučných statích („jazykové koutky“ apod.). K těm všem, byť v nutné selekci, rovněž přihlížíme, i zde se snahou o kontrastivní přístup: je-li to možné, jsou lotyšské zdroje konfrontovány s jejich českými protějšky. Tak jsou např. údaje z nejnovější lotyšské mluvnice (MLG) soustavně zanášeny do naší databáze (atribut Gramatika elementu Lemma, viz 2.4.3.15), srovnávány s popisem v předchozí lotyšské gramatice (MLLVG I, II) a s jinými poznatky z odborné lotyšské literatury, ale potenciálně i té české. Opouští se tak úzce lexikografická perspektiva a posiluje se komplexní charakter znalostní databáze lotyštiny (viz též závěr kap. 2.4).

### 2.3.4 Internetové zdroje

Z internetových pramenů byly vybrány jen ty dostatečně hodnověrné, a to zvláště k ověřování terminologie: např. otevřená encyklopedie Wikipedia,<sup>61</sup> ter-

---

<sup>60</sup> V ČLN je číslicí 1 v horním indexu před lemmatem (<sup>1</sup>ulicje f, (-e, Gpl -) iel/a f [...]) označena příslušnost lexému k první tisícovce frekvenčního žebříčku podle korpusu SYN2000. Podobně by šlo frekvenční špičku vyznačit i v hesláři LČ, bylo by to však záhodno učinit na základě dat z mnohem většího korpusu, než je LVK. Srov. návrh rozlišování různých frekvenčních pásem v Čermák 1995c: 100)

<sup>61</sup> Česká wikipedie má aktuálně 348 782 článků (je tak 26. nejobsáhlejší jazykovou mutací této encyklopedie), lotyšská 66 781 článků (67. v pořadí). Data jsou aktuální k 20. 3. 2016, citováno podle [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias).

minologická databáze *EuroTermBank*,<sup>62</sup> *AkadTerm*<sup>63</sup> a *Termnet.lv*<sup>64</sup> encyklopedie rostlinných a živočišných druhů vyskytujících se v Lotyšsku *Latvijas daba*<sup>65</sup> ve srovnání s podobnou, leč šířeji koncipovanou českou databází *BioLib*,<sup>66</sup> portál *letonika.lv*, zastřešující 11 encyklopedií, 13 slovníků a další audiovizuální obsah, archiv lotyšského tisku *news.lv* aj. Mimoto lze internet využít k bohaté excerpci (viz následující bod), vždy však s přihlédnutím k důvěryhodnosti pramenu, typu textu, komunikační situaci, případně i dalším aspektům. Obzvláště problematické jsou frekvenční údaje poskytované nejrozličnějšími internetovými vyhledávači; nekritické vyvozování soudů na základě takovýchto dat je velmi ošemetné (Svensén 2009: 60, Čermák 2007: 68).

### 2.3.5 Excerpta

Doplňkovým pramenem, dnes značně upozaděným korpusovými daty, jsou excerpta jak z psaných, tak mluvených textů. Oproti dřívějšímu má dnešní lexikograf přístup k rozličným typům textů na internetu (diskusní fóra, blogy, sociální sítě, internetová média, kanál YouTube aj.), a tedy i k nepřehlednému množství dat, navíc jsou tu stále tištěné zdroje, především společenská periodika, ideální pro excerpci vzhledem ke své tematické šíři a pestrosti, ale též pro tíhnutí k mluvené podobě jazyka (pronikání hovorových i nespisovných forem) a k determinologizaci (přecházení slov z úzkého oborového milieu do celospolečenského oběhu). Odborná a profesní periodika a monografie lze zase využít k cílenému doplnění chybějící oborové terminologie, popřípadě k jejímu ověření. Výběr excerptovaného materiálu bývá spíše nahodilý a nutně subjektivní; ex-

---

<sup>62</sup> Hlavním koordinátorem projektu ETB je lotyšská softwarová společnost Tilde, věnující se vývoji a aplikaci jazykových technologií; na projektu se podílí rovněž lotyšská Akademie věd a dalších sedm subjektů z Litvy, Estonska, Polska, Maďarska, Německa a Dánska. Jeho primárním cílem je „harmonizovat a konsolidovat práci terminologů v nových členských státech EU, přenášet zkušenosti z jiných terminologických sítí EU a integrovat kompetence a snahy nově přistoupivších zemí“ (<http://www.eurotermbank.com/about.aspx>).

<sup>63</sup> <http://termini.lza.lv/term.php>. Aktuálně obsahuje přes milion termínů v šesti jazycích (lotyština, angličtina, ruština, němčina, francouzština, latina).

<sup>64</sup> <http://termnet.lv/>. Aktualizace databáze skončila v roce 2009, obsahuje necelých 150 tisíc hesel v týchž šesti jazycích jako *AkadTerm*.

<sup>65</sup> <http://www.latvijasdaba.lv/>. Aktuálně s informacemi o téměř 5500 biologických druzích.

<sup>66</sup> <http://www.biolib.cz/>. Aktuální statistika uvádí 5275 slovníkových hesel.

cerpující osoba má tendenci všímat si jevů neobvyklých, zatímco jazykové korpusy odhalují typický úzus,<sup>67</sup> jakkoliv obě metody akvizice dat zaručují skutečně autentické ukázky parole, nikoliv jen umělé konstrukty, obsažené často ve starších slovnících. Lze si nicméně představit účelovou excerpci např. kolekce jídelních lístků, z nichž by (v případě česko-cizojazyčného překladového slovní-slovníku) lexikograf vytěžil některá „specifika“ české gastronomie, která by mu z jiných zdrojů mohla uniknout (lemmata *utopenec*,<sup>68</sup> kolokace *nakládany hermelín*, *španělský ptáček*<sup>69</sup> apod.). Odvažujeme se proto tvrdit, že ani dnes není tradiční praxe excerpování překonána, třebaže byla její role jazykovými korpusy zastíněna, a může posloužit coby korektiv slovníkářské práce.

Obhlédneme-li tedy svou materiálovou základnu z celkové perspektivy (souhrnný přehled zdrojů využívaných při tvorbě LČ je uveden v bibliografii k této práci), lze oprávněně tvrdit, že je dostatečně velkého množství i pestrého složení<sup>70</sup> a dovoluje tak vytvořit kvalitní lexikografické dílo. Ve skutečnosti je potenciálních podkladů pro práci na slovníku k dispozici více než kdykoliv dříve, až hrozí riziko, že v takovém mnohohlasí mohou být skutečně zásadní zdroje nedoceny na úkor těch marginálních. Prevenci vidíme v jasné hierarchizaci pramenů, respektive informací, které z nich do databáze zaznamenáváme: zdaleka ne všechny totiž budou v LČ použity. Klíčové jsou pro nás zdroje z prvních tří skupin, tj. korpusová data a údaje čerpané z tištěných a elektronických slov-

---

<sup>67</sup> „Individuals notice the unusual, whereas corpora will highlight the common and frequent senses.“ (Varantola 2003: 232) Srov. též Cvrček–Kovaříková 2011: 120 či Svensén 2009: 44–45.

<sup>68</sup> Slovo *utopenec* v SSČ (i SSJČ) má jediný význam: ‚kdo se utopil, utopený‘, srov. 87 výskytů lemmatu *utopenec* v korpusu SYN2015, z nichž 51 má původní význam, 30 označuje pokrm, 6 zbývajících (výhradně v beletristických textech, najmě poezii) je sporných, neboť záměrně využívá polysémii daného slova. Viz [http://ps://kontext.korpus.cz/view?q=~BuCTzIrr&attr\\_allpos=kw&attrs=word%2Clemma&corpname=syn2015&ctxattrs=word&pagesize=40&refs=%3Ddoc.title%2C%3Ddoc.author%2C%3Ddoc.publisher%2C%3Ddoc.pubyear&structs=p&viewmode=kwic&](http://ps://kontext.korpus.cz/view?q=~BuCTzIrr&attr_allpos=kw&attrs=word%2Clemma&corpname=syn2015&ctxattrs=word&pagesize=40&refs=%3Ddoc.title%2C%3Ddoc.author%2C%3Ddoc.publisher%2C%3Ddoc.pubyear&structs=p&viewmode=kwic&)

Oba významy správně rozlišuje ČLN (317).

<sup>69</sup> Tuto kolokaci zná jak SSČ, tak i SSJČ, na rozdíl od kolokace *nakládany hermelín*, což je jistě podmíněno dobou vzniku těchto slovníků, stejně je tomu tak ale i v ČLN z roku 2006; srov. SYN2015: *španělský ptáček* 22, *nakládany hermelín* 6, *naložený hermelín* 1.

<sup>70</sup> Srov. nedávnou výchozí situaci autorky ČLN, jež neměla k dispozici ani lotyšský, ani paralelní česko-lotyšský korpus (Nikuļceva 2013: 214).

níkových a encyklopedických příruček a také odborné jazykovědné literatury; s internetovými daty je třeba nakládat obezřetně, rovněž tak nelze přeceňovat ani roli excerpt. V ideálním případě by všech pět typů pramenných materiálů mělo být vzájemně konfrontováno, interagovat mezi sebou a doplňovat se (Svensén 2009: 62–63); výsledná podoba hesla by měla do určité míry zohlednit každý z nich. Neméně důležitá je i aktuálnost zdrojů, o něž se opíráme: jen s nimi lze podat autentický obraz současné lotyštiny.

## 2.4 Technický aspekt tvorby LČ

LČ je sestavován na základě lexikální databáze, vznikající ve specializovaném lexikografickém softwaru TshwaneLex (dále TLex). Tento program, vyvíjený jihoafrickým programátorem Davidem Joffem od roku 2003, byl použit při tvorbě již mnoha moderních slovníků<sup>71</sup> a odbornou veřejností bývá oceňován pro svou univerzálnost: autorovo tvrzení, že jde o nástroj použitelný nezávisle na jazyku (kombinaci jazyků), plně přizpůsobitelný uživatelským potřebám (včetně možnosti kompletní lokalizace do libovolného jazyka) a všestranně uživatelsky přívětivý, není pouhou reklamní frází. Podstatné je rovněž hledisko ekonomické: TLex zajišťuje vedle výrazné úspory časové také snížení výrobních nákladů zefektivněním lexikografické práce.

V dnešní době se speciální lexikografické programy při tvorbě slovníků stále více prosazují na úkor nástrojů dřívějších, např. textových editorů, zpracovávajících veškerý jazykový obsah lineárně. Práce v nich byla velmi namáhavá a vinou nepřítomné hierarchické organizace vznikajícího díla též nepřehledná a těžkopádná. Navíc ani přes veškerou snahu nebylo možno zabránit poměrně častým chybám zaviněným lidským faktorem (nesprávné abecední řazení, ne-

---

<sup>71</sup> Na internetových stránkách tohoto softwarového nástroje jsou jako příklad uváděny zhruba čtyři desítky slovníků všech typů a jiných lexikografických příruček, jedná se přitom o pouhý výběr. Viz <http://tshwanedje.com/dictionaries/>.

jednotné formátování, nejednotný metajazyk, nefunkční odkazy).<sup>72</sup> Pochybení tohoto typu TLex eliminuje a na eventuální diskrepance uživatele upozorní chybovým hlášením. Případné změny v nastavení struktury hesel se projeví okamžitě v celém souboru.<sup>73</sup> Velice účinným a užitečným prostředkem je rovněž možnost rychlého a snadného vyhledávání a filtrace dat; pomocí filtrů lze též snadno korigovat požadovaný objem slovníku a jednou zaznamenaná data opakovaně využít k různým účelům (srov. Čermák 1995c: 97, Vondříčka 2011: iii, 14, 17, Atkins–Rundell 2008: 100, Bergenholtz–Tarp 1995: 40 aj.). – Obecně lze tvrdit, že lexikografický software přispívá ke konzistenci a koherenci výsledného díla a potlačuje negativní vliv lidského faktoru.

Lexikální databáze v TLexu je dokumentem formátu XML (*eXtensible Markup Language*, čes. rozšířitelný značkovací jazyk), jehož struktura je specifikována pomocí tzv. definice typu dokumentu<sup>74</sup> (*Document Type Definition*, DTD). Definovány a pravidly vymezeny jsou typy elementů (*elements*), tvořících hierarchickou, stromovou<sup>75</sup> strukturu slovníkového hesla, včetně jejich vzájemných vztahů (rodič–potomek [*parent–child*]), a jejich atributy (*attributes*), tj. vlastní data představující náplň těchto systémových jednotek. Při zakládání nového lexikografického projektu nabídne TLex přednastavenou, defaultní DTD šablonu, jež z pochopitelných důvodů, tkvících ve specifitě každého slovníku, nemůže být použita bez náležitých úprav. Také v našem případě byla notně přepracována (především na úrovni atributů) tak, aby vyhovovala našim potře-

---

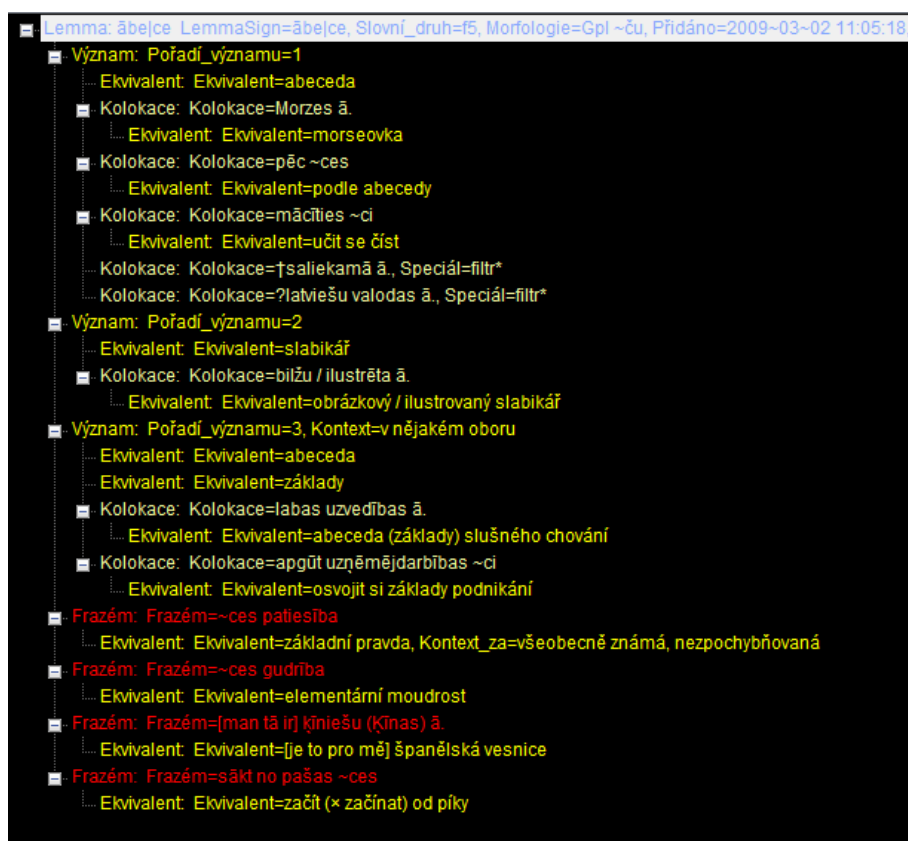
<sup>72</sup> K tomu přiléhavě Ridings (2003: 212): „Jsou to úkoly buď pro pedanta, nebo pro software.“

<sup>73</sup> Dřívější obavy, vyjádřené např. Zgustou (1971: 352), jenž o změnách ve struktuře hesel či metajazyku v pokročilejších fázích lexikografické práce hovoří jako o „pohromě“ a „opravdových katastrofách“ (srov. též Landau 1991: 83), se pak z dnešní perspektivy jeví jako přehnané. Avšak i v TLexu přetrvává riziko např. nejednotnosti metajazyka (glosy typu *o muži* × *muž*, *apod.* × *ap.*). Těmto potenciálním problémům je proto třeba věnovat zvýšenou pozornost: jednak při počáteční, přípravné fázi, neméně i ve fázi dokončovací (revize, korektury).

<sup>74</sup> Též se hovoří o gramatice slovníku (*dictionary grammar*), srov. „bezkontextová gramatika popisující hnízdování elementů“ (Čermák 1995b: 60) nebo jiná pojetí: „sada pravidel“, „soubor podmínek“ (obě Pořízka 2014: 82) či „strukturní šablona“ (<https://wiki.korpus.cz/doku.php?id=pojmy:parser>).

<sup>75</sup> Často se v tomto případě pro lepší názornost užívá přímo metafory stromu: jejím prizmatem lze na elementy nahlížet jako na větve, na atributy pak jako na listy (MTL: 60).

bám a požadavkům.<sup>76</sup> V následující podkapitole bude uveden přehled elementů a jejich atributů využívaných pro potřeby chystaného LČ, s upozorněním na změny oproti defaultní šabloně.<sup>77</sup> Naopak čistě formální rysy databáze, odrážející uživatelské osobní preference (fonty, barvy, klávesové zkratky a jiná uživatelská nastavení), jsou pominuty jakožto irelevantní pro výsledný vzhled slovníku; grafická podoba vzorových hesel předkládaných v této práci není definitivní a může v průběhu příprav slovníku doznat změn (byť ne radikálních), vždy však musí být brán zřetel na uživatele a jeho co nejsnazší a nejpohodlnější orientaci.



Obrázek č. 5: Stromová struktura hesla *ābece* v TLexu, srov. níže výstup ve formátu RTF

<sup>76</sup> K výše zmíněným přednostem TLexu je dobré zmínit další vstřícný krok vůči uživateli, jenž dokáže původní DTD strukturu upravit či předdefinovat i bez znalosti jazyka XML, tudíž sám, bez pomoci odborníka v oblasti IT. (MTL: 60)

<sup>77</sup> Viz též popis defaultně nastavených prvků a atributů v manuálu TLexu (MTL: 61–63).

**ābe|ce** f<sub>5</sub> (Gpl ~ču) **1.** *abeceda* • Morzes ā. *morseovka* • pēc ~ces *podle abecedy* • mācīties ~ci *učit se číst* • <sup>†?</sup>saliekamā ā. • <sup>?</sup>latviešu valodas ā.<sup>78</sup> || **2.** *slabikāř* • bilžu / ilustrēta ā. *obrázkový / ilustrovaný slabikář* || **3.** (v nějakém oboru) *abeceda, základy* • labas uzvedības ā. *abeceda (základy) slušného chování* • apgūt uzņēmējdarbības ~ci *osvojit si základy podnikání* || ♦ ~ces patiesība *základní pravda* (všeobecně známá, nezpochybňovaná) ♦ ~ces gudrība *elementární moudrost* ♦ [man tā ir] ķīniešu (Ķīnas) ā. *[je to pro mě] španělská vesnice* ♦ sākt no pašas ~ces *začít* (× *začínat*) *od píky*

Předem avizovaného přehledu elementů je třeba upozornit na jejich dvojí povahu, respektive rozlišovat elementy interní, předprogramované (*Tshwane-Lex-internal*, MTL: 66), a tudíž neměnné (pouze přejmenovatelné), zajišťující „pokročilejší funkčnost“ (ibid.) databáze, od těch ostatních, jejichž správa spadá plně do uživatelské kompetence (viz též třída *User* těchto elementů v DTD struktuře TLexu<sup>79</sup>). Interní elementy jsou v DTD struktuře kromě specifické třídy (*class*) odlišeny rovněž graficky, šedou barvou namísto standardní černé. Jedná se o tyto elementy: Slovník, Jazyk, Lemma a Reference. V rámci elementů neinterních dále rozlišujeme elementy veřejné od neveřejných: neveřejné jsou nositeli informací zaznamenaných v databázi, přičemž s jejich začleněním do LČ se nepočítá. Status jednotlivých elementů (i jejich atributů) lze snadno a pohodlně měnit zaškrtnutím či odškrtnutím okénka *Visible (as child of selected element)* v DTD struktuře – o tom, které informace budou zpřístupněny uživatelům LČ a které jim zůstanou skryty, se tak dá rozhodnout až v úplném závěru příprav slovníku.

#### 2.4.1 Element Slovník (defaultně *Dictionary*, třída *DocumentRoot*)

je interním, kořenovým (*top-level*, MTL: 67) elementem celého dokumentu, obsahujícím v případě LČ dva elementy Jazyk. Jeho atributy (*Description*, *CreatedBy*, *Copyright*, *URL*, *SoundPath*) jsou vesměs interní, sloužící pro po-

<sup>78</sup> Poslední dvě kolokace jsou ponechány bez překladu záměrně: jejich zařazení do LČ je nepravděpodobné, zůstanou tak nejspíš, s příslušnou distinkční značkou (v horním indexu před kolokací), pouze v databázi. V prvním případě jde o realii zastaralou (skládací abeceda [školní pomůcka]), doloženou v LVTK v jediném výskytu, v druhém případě („lotyšská abeceda“) o spojení hypoteticky použitelné, avšak v korpusech nedoložené (běžně se v tomto smyslu používá synonyma *alfabēts*).

<sup>79</sup> Nabídka *Dictionary – Customise DTD (dictionary grammar)*.



třeby samotné databáze, a uživatelem neměnné, nebudeme se jimi zde proto zabývat.

#### 2.4.2 Element Jazyk (defaultně *Language*, třída *Section/Language*)

je interním potomkem elementu Slovník a kořenovým elementem z hlediska makrostruktury slovníku: slouží jako rodič elementů Lemma. Atributy elementu Jazyk (*Options*, *AlphabetQuicklinks*, *SearchScript*, *ClickScript*, *xml:lang*) jsou stejně jako atributy elementu Slovník čistě interní, uživatelsky neměnné.

#### 2.4.3 Element Lemma (defaultně *Lemma*, třída *Entry*)

je interním potomkem elementu Jazyk a kořenovým elementem z hlediska mikrostruktury slovníku: obsahuje vlastní slovníkové heslo se všemi informacemi o něm. Slouží jako rodič všech níže uvedených elementů. Mimoto zahrnuje v případě naší databáze tyto atributy:

##### 2.4.3.1 Atribut Incomplete

je interním atributem nabývajícím buď kladnou (heslo není dokončeno), nebo zápornou (heslo je dokončeno) hodnotu.

##### 2.4.3.2 Atribut LemmaSign

je interním atributem obsahujícím samotné slovníkové heslo. Využívá se k automatickému řazení hesel v hesláři, není proto možné jej odstranit ani změnit.

##### 2.4.3.3 Atribut Homonymum

slouží k rozlišení jednotlivých homonym.<sup>80</sup> TLex v tomto ohledu nabízí širokou řadu možností, v LČ bude použito římských číslic v horním indexu bezprostředně následujícím příslušné lemma ( $t\bar{a}^1$ ,  $t\bar{a}^2$ ,  $t\bar{a}^3$ ).

---

<sup>80</sup> Z hlediska TLexu jsou za homonyma považovány pouze případy, kdy „více hesel sdílí stejnou pravopisnou formu lemmatu (tj. hesla mají stejný pravopis)“ (MTL: 62), což v našem případě může představovat problém. Součástí lemmat ohebných slovních druhů bývá také svislice (kolmička), sloužící k označení místa, od něhož se daný výraz ohýbá. Pro TLex tak např. nejsou homonymní dvojice slov *takt/s* („takt, ohleduplnost“) a *tak/ts* („takt

#### 2.4.3.4 Atribut Fonologie

slouží k fonologické transkripci daného lemmatu, respektive jeho relevantní části, případně k zachycení nepravidelností a jiných specifických případů. Fonologický přepis, přizpůsobený potřebám českého uživatele, následuje bezprostředně po lemmatu, je tištěn stojatým netučným písmem v hranatých závorkách.

Více o fonologickém popisu lexémů v LČ viz 3.2.2.

#### 2.4.3.5 Atribut Slovní\_druh

označuje příslušnost daného lemmatu k určité třídě slov, především jeho slovnědruhovou platnost. Slovnědruhové markery mají podobu zkratk vycházejících z latinských názvů jednotlivých slovních druhů a píší se menším stojatým písmem, s případným blíže specifikujícím příznakem v dolním indexu (deklinční třídy u substantiv, konjugační třídy u verb aj.).<sup>81</sup> Vybrat je lze z libovolně upravovatelného seznamu zaškrtnutím příslušného okénka (*Data type: Select from list /,multiple“/*). Naopak u zjevných případů (hesla morfémové povahy, zkratky, části složenin aj.) lze slovnědruhovou zkratku vypustit. Jednotlivé slovní druhy spolu s příslušnými paradigmaty budou uživateli blíže představeny v gramatickém přehledu následujícím po hesláři.

#### 2.4.3.6 Atribut Morfologie

slouží k záznamu morfologických údajů o lemmatu (odlišná morfologie pro jednotlivé významy je řešena na úrovni elementu Význam jakožto potomka elementu Lemma), zejména takových, jež dané lemma jednoznačně přiřazují k příslušnému flektivnímu paradigmatu, a tvarů nepravidelných či těch, které by českému uživateli mohly činit potíže (u takových se navíc počítá s odkazem na příslušnou pasáž gramatického popisu na konci LČ). Tvary jsou ve snaze ušetřit

---

[rytmický celek]‘) nebo *pa/ts* (‘sám, jediný‘) a *pat/s* (‘pat‘). Řešení vidíme v dodatečné editaci rukopisu slovníku v textovém editoru, kam bude vyexportován.

<sup>81</sup> Tento příznak postihuje formální rysy lexému, kdežto slovnědruhová značka se opírá především o sémantiku hesla. Např. zkratka *num<sub>adj</sub>* označuje číslovky skloňované podle složené adjektivní deklinace, zatímco *num<sub>m1</sub>* číslovky skloňované podle 1. (maskuliní) deklinační třídy. První třída lexémů vykazuje formální podobnost s adjektivy (řadové číslovky), druhá se substantivy (číslovky substantivní povahy: *milion, miliarda* apod.).

místo uváděny ve zkrácené podobě: neměnnou část hesla (u lemmatu nalevo od svislice | ) zastupuje znak vlnovky ~, po němž následuje část proměnlivá. Od tohoto zkráceného zápisu upouštíme v případech, kdy by svislice oddělovala jediné písmeno (např. *āzis*, *āža* namísto *ā|zis*, *~ža*), či u tvarů supletivních. Morfologické údaje následují po slovnědruhovém zkratkce, jsou od ní odděleny kulatými závorkami a tištěny menším stojatým písmem.

Více o gramatickém popisu lexémů v LČ viz 3.2.3.

#### 2.4.3.7 Atribut Valence

slouží k postižení valence daného lemmatu v širokém slova smyslu, tištěn je menším kurzivním písmem v kulatých závorkách.

#### 2.4.3.8 Atributy Přidáno a Změněno

jsou interními atributy sloužícími k snazší orientaci při organizaci práce: obsahují časové údaje o vytvoření, respektive poslední změně daného hesla.

#### 2.4.3.9 Atribut Zdroj

je neveřejným atributem sloužícím k zaznamenání zdrojů, v nichž dané heslo figuruje, ať už se jedná o slovníky tištěné či elektronické, případně jiné prameny (internet, ruční excerpta apod.).

#### 2.4.3.10 Atribut Stylový\_příznak

charakterizuje lemma zejména z hlediska diastratického. Stylistické markery mají podobu zkratk (bez tečky) menším stojatým písmem ve špičatých závorkách (např. <hanl>, <expr>), vybírají se z libovolně upravovatelného seznamu zaškrtnutím okénka (okének) s příslušnou charakteristikou (*Data type: Select from list /multiple/*).

Více k lingvistickým aspektům stylistického značení viz 3.2.6.

#### 2.4.3.11 Atribut Oborový\_příznak

slouží k postižení diatechnické roviny, tedy primárně k vydělení odborné terminologie jakožto specifické vrstvy lexikonu. Oborové markery se označují rovněž zkratkami menším stojatým písmem ve špičatých závorkách, od markerů stylových je lze snáz odlišit podtržením (např. <biol>, <lingv>). Vybírají se rovněž ze

seznamu zaškrtnutím okénka (okének) s názvem příslušného vědního oboru či disciplíny. V případě některých oborů je odborný termín doplněn svým latinským protějškem (viz atribut Latina).

Více k lingvistickým aspektům oborového značení viz 3.2.6.

#### 2.4.3.12 Atribut Dešifrace\_zkratky

je veřejným atributem pro zkratky a zkratková slova, obsahujícím jejich plnou podobu. Jakkoliv jde o specifickou skupinu hesel, v LČ budou, oproti starší lexikografické praxi, zahrnuta v obecném hesláři.

Všechny následující atributy jsou neveřejné povahy, slouží pouze osobním potřebám sestavovatele databáze a s publikováním informací v nich uvedených se ve výsledné podobě LČ nepočítá.

#### 2.4.3.13 Atribut Speciální\_značky

je neveřejným atributem sloužícím k záznamu dalších doplňkových informací potřebných pro tvůrce databáze („problematická“ hesla, nutnost konzultace s rodilým mluvčím, vydělení specifických reálií či tzv. falešných přátel apod.).

#### 2.4.3.14 Atribut Poznámky

je neveřejným atributem obsahujícím zpracovatelův poznámkový aparát.<sup>82</sup>

#### 2.4.3.15 Atribut Gramatika

je neveřejným atributem sloužícím k evidenci odkazů na dvojici lotyšských mluvnic (LVG; MLLVG I, II; ke kontrastivní analýze však potenciálně i těch českých) a další odborné zdroje, kde se o daném lexému pojednává.

#### 2.4.3.16 Atribut Etymologie

je neveřejným atributem postihujícím u vybraných hesel jejich původ, případně též jejich diachronní podobu.

#### 2.4.3.17 Atribut Exil

---

<sup>82</sup> Tímto atributem, zůstane-li neveřejným, je možno (a snad i záhodno) vybavit každý element; u sémanticky komplikovaných hesel může vzniknout potřeba opatřit poznámkou např. jednu konkrétní kolokaci či frazém. V dalším popisu však od této možnosti upouštíme.

zaznamenává pravopisnou variantu užívanou lotyšskými emigranty. O exilové varietě lotyštiny bude stručná zmínka v gramatickém přehledu na konci LČ, v samotném hesláři však od ní bude odhlíženo. Na rozdíl od veřejného elementu Varianta je atribut Exil neveřejný a slouží interním potřebám zpracovatele.<sup>83</sup>

#### 2.4.3.18 Atribut Korpus

je neveřejným atributem sloužícím k záznamu údajů o četnosti daného lemmatu v korpusech. Defaultně nastavený atribut *Frequency* vyžadoval hodnotu celého čísla (*Data type: Integer*), což znemožňovalo zápis více údajů a také použití zkratk pro rozlišení jednotlivých korpusů (např. SYN2010: 225).

#### 2.4.3.19 Atribut Synonyma

je neveřejným atributem obsahujícím synonyma (případně odrážejícím i další paradigmatické vztahy) daného lemmatu.

#### 2.4.3.20 Atribut Ref

je neveřejným atributem odkazujícím na jiné heslo (jiná hesla) ve slovníku. Na rozdíl od veřejného elementu Reference slouží interním potřebám zpracovatele.

### 2.4.4 Element Varianta (nově přidaný element, třída *User*)

slouží k záznamu ortografické či morfologické varianty lemmatu (méně frekventované, příznakové aj. dublety), následují-li tyto v abecedě bezprostředně po sobě. Je to přímý potomek elementu Lemma a má stejné atributy jako on (viz jejich popis výše). Oproti dílčímu atributu Varianta elementů Kolokace a Frázém má komplexnější povahu: funguje jako samostatný element s řadou atributů.

---

<sup>83</sup> Obdobně lze zavést nové atributy, jež by sloužily k zachycení jiných variant lexému: diachronní (historické formy mohou být těž zaznamenávány v rámci atributu Etymologie), dialektální, eventuálně i jiné.

#### 2.4.5 Element Význam (defaultně *Sense*, třída *User*)

je elementem nesoucím jednotlivé významy daného lexému. Funguje jako potomek elementů Lemma, Kolokace a Frazém a jako rodič elementů Ekvivalent, Kolokace, Frazém, Reference, Excerptum a Ref. Zahrnuje následující atributy:

##### 2.4.5.1 Atribut Pořadí

slouží k oddělení jednotlivých významů polysémních lexémů. V LČ bude k tomuto účelu využito arabských číslic, tištěných tučně a následovaných tečkou: **1.**, **2.**, **3.** Jednotlivé významy jsou od sebe pro přehlednost odděleny dvojí-  
tou svislicí (||) a zpravidla též arabskou číslicí; jde-li o drobnější sémantickou nuanci či dílčí aspekty téhož významu, nahrazuje znak dvojité svislice středník užívaný v jiných slovnících k označení této distinkce.

##### 2.4.5.2 Atribut Pořadí\_alter

slouží k alternativnímu řazení významů na úrovni víceslovných lexémů. Odlišné významy tu budou označeny malými písmeny s kulatou půlzávorkou tištěnými tučným kurzivním písmem: ***a)***, ***b)***, ***c)***.

##### 2.4.5.3 Atribut Slovní\_druh

slouží k subkategorizaci dílčího významu daného lemmatu z hlediska slovnědruhového úzu.

##### 2.4.5.4 Atribut Morfologie

slouží k záznamu morfologických údajů o dílčím významu daného lemmatu.

##### 2.4.5.5 Atribut Valence

slouží k postžení valence dílčího významu daného lemmatu.

##### 2.4.5.6 Atribut Kontext

slouží ke kontextovému upřesnění dílčího významu daného lemmatu.

##### 2.4.5.7 Atribut Funkce

slouží pro popis funkce u gramatických slov.

##### 2.4.5.8 Atributy Stylový\_příznak a Oborový\_příznak

jsou paralelami stejnojmenných atributů elementu Lemma, omezujícími se však nikoliv na celé lemma, ale na jeho dílčí význam.

#### 2.4.6 Element Ekvivalent (defaultně *TE*,<sup>84</sup> třída *User*)

je potomkem elementů Význam, Kolokace a Frazém a může sloužit jako rodič elementu Excerptum a Ekvivalent.<sup>85</sup> Původní řešení, kdy ekvivalent nebyl samostatným elementem, ale pouze atributem elementu Význam (resp. Kolokace a Frazém), se neosvědčilo: takovýto atribut nutně vyžadoval velké množství doplňkových atributů sloužících k jeho komplexní charakteristice (k záznamu jeho výslovnosti, slovnědruhovému příslušnosti, valence, kontextového upřesnění aj.), a toto nemalé množství se navíc znásobovalo (potenciálně vysokým) počtem ekvivalentů. Databáze se tak rozrostla o četné kolonky, které však zůstávaly začasné nevyplněny, což vedlo ke ztrátě přehlednosti. Rovněž docházelo k prolínání pravé a levé strany slovníku, které by měly být zřetelně odděleny. Z tohoto důvodu považujeme stávající řešení, tj. status ekvivalentu jakožto samostatného elementu, jež lze podle potřeb přidávat v libovolném množství a na příslušné místo v hierarchické struktuře databáze, za pružnější a efektivnější.

Element Ekvivalent zahrnuje následující atributy:

##### 2.4.6.1 Atribut Ekvivalent

obsahuje vlastní český ekvivalent lotyšského lexému (lemmatu, kolokace, frazému). Je tištěn kurzivním písmem, což je primární prostředek pro oddělení české části od lotyšské.

---

<sup>84</sup> *Translation Equivalent*.

<sup>85</sup> K tomuto řešení se uchylujeme v případě vidových protějšků: ty musí mít rovněž povahu elementu, nikoliv atributu proto, aby je bylo v budoucnu případně možno doplnit o komplexní charakteristiku; rovněž tak může mít ekvivalent více než jeden vidový protějšek. Vzniká tak struktura Lemma–Význam–Ekvivalent–Ekvivalent (např. *atbraukt–l.–přijet–přijíždět*).

#### 2.4.6.2 Atribut Latina

obsahuje latinský (případně jiný cizojazyčný<sup>86</sup>) ekvivalent lotyšského lexému a jeho českého protějšku, především u zoologických a botanických termínů. Lze tu však rovněž uvádět ekvivalentní symboly a značky (např. chemické prvky, hudební značky, interpunkce, písmena řecké abecedy aj.). Takovýto doplňující ekvivalent má sloužit k snazší a jednoznačné identifikaci daného lexému. Je tištěn menším stojatým písmem v kulatých závorkách.

#### 2.4.6.3 Atribut Valence

postihuje valenci českého ekvivalentu. Stejně jako paralelní atribut elementu Lemma, respektive Význam je tištěn menším kurzivním písmem v kulatých závorkách.

#### 2.4.6.4 Atributy Kontext\_před a Kontext\_za

slouží k bližšímu kontextovému upřesnění daného ekvivalentu, ať už formou synonymního prostředku, kolokační glosy, encyklopedické informace či jinými prostředky, a to podle potřeby buď před ekvivalentem, nebo za ním. Tisknou se rovněž menším stojatým písmem v kulatých závorkách.

#### 2.4.6.5 Atribut Fonologie<sup>87</sup>

zachycuje výslovnost u vybraných ekvivalentů, které by mohly činit potíže lotyšskému uživateli (slova cizího původu, dvojí možná výslovnost skupin *di*, *ti*, *ni*, zkratky aj.). Tiskne se menším stojatým písmem v hranatých závorkách.

#### 2.4.6.6 Atribut Morfologie

popisuje vybrané ekvivalenty z hlediska gramatického; je určen výhradně lotyšskému uživateli, jemuž pomáhá přiřadit daný ekvivalent k příslušné třídě slov a paradigmatu. U substantiv bude uváděn rod (u maskulin pak životnost/neživotnost) a další specifika (singularia/pluralia tantum aj.), u verb pak

---

<sup>86</sup> Týká se to především cizojazyčných zeměpisných jmen, srov. lot. *Vīne*, čes. *Vídeň*, něm. *Wien* či lot./čes. *Varšava*, pol. *Warszawa*. Daný atribut se však osvědčil i na úrovni kolokací: např. *Centrāleiropas laiks* (‘středoevropský čas’), kdy atribut Latina obsahuje dvojí zkratku: *SEČ*, *CET*.

<sup>87</sup> Atributy 2.4.6.5–2.4.6.7 jsou součástí DTD struktury, v tuto chvíli je však ponecháváme prázdné. Dodatečné informace o českých ekvivalentech, určené pro potenciální lotyšské uživatele LČ, budou zpracovány samostatně, v pozdější fázi, a nejsou proto ani součástí ukázkových hesel uváděných v této práci.



vid; s dalšími morfologickými údaji se vzhledem k omezenému prostoru nepočítá, komplexní morfologickou charakteristikou budou vybaveny pouze lotyšské lexémy. Slovnědruhovná zkratka je tištěna malým stojatým písmem v dolním indexu bezprostředně za ekvivalentem, v případě víceslovných ekvivalentů za klíčovým slovem (substantivum, verbum).

#### 2.4.6.7 Atribut Stylový\_příznak

slouží k postižení odlišné stylistické charakteristiky českého ekvivalentu od lotyšského lexému. Mělo by však jít o případy výjimečné, v zásadě je třeba hledat takový ekvivalent, jenž lotyšskému lexému odpovídá v obou složkách významů: nociónální i pragmatické; na případné neshody je třeba explicitně upozornit (glosou, poznámkou o úzu apod.). Podobně jako stejnojmenný atribut elementu Lemma, respektive Význam je tento atribut tištěn menším stojatým písmem ve špičatých závorkách, avšak jakožto doplňková informace určená primárně lotyšskému uživateli se nachází za ekvivalentem v dolním indexu a má lotyšskou variantu příslušné stylistické zkratky.

#### 2.4.6.8 Atribut Poznámka

je interním, neveřejným atributem sloužícím k případným poznámkám o daném ekvivalentu.

#### 2.4.7 Element Kolokace (defaultně *Subentry*,<sup>88</sup> třída *User*)

je potomkem elementů Lemma a Význam a rodičem elementů Ekvivalent, Význam, Reference a Excerptum. Slouží k exemplifikaci daného lexému a k naznačení jeho základní kolokability pomocí syntagmat s různou mírou ustálenosti. V případě lexémů s omezenou kolokabilitou se tento element napojuje přímo na element Lemma a český ekvivalent je hledán k příslušné kolokaci, nikoliv k lemmatu (či jeho dílčímu významu). Element Kolokace zahrnuje následující atributy:

---

<sup>88</sup> Ve starších verzích TLexu *Combination*.

#### 2.4.7.1 Atribut Kolokace

obsahuje vlastní kolokaci, tedy víceslovné pojmenování, ať už volné či ustálené. Kolokace jsou tištěny polotučným stojatým písmem a pro snazší orientaci v hesle je každá uvozena speciálním znakem kolečka •.

#### 2.4.7.2 Atribut Valence

postihuje valenci uvedené kolokace. Je, stejně jako veškeré údaje o valenci, tištěn menším ležatým písmem v kulatých závorkách.

#### 2.4.7.3 Atribut Varianta

slouží k zachycení méně časté či příznakové varianty dané kolokace. Uvádí se bezprostředně za kolokací v kulatých závorkách obyčejným písmem.

#### 2.4.7.4 Atributy Stylový\_příznak a Oborový\_příznak

jsou paralelami stejnojmenných atributů elementu Lemma, slouží k diasystematickému omezení kolokace, v níž dané lemma figuruje. Způsob zápisu je stejný jako u všech stylistických/oborových markerů: menším stojatým písmem (podtrženým v případě oborových značek) ve špičatých závorkách.

#### 2.4.8 Element Frazém (nově přidáný element, třída *User*)

je potomkem elementů Lemma a rodičem elementů Ekvivalent, Význam, Reference a Excerptum. Slouží k záznamu lexémů frazeologické povahy: frazémů, přísloví, okřídlených rčení, citátů apod. Struktura atributů je totožná s atributy předchozího elementu Kolokace (z hlediska čistě formálního jde v obou případech o kombinace dvou a více slov): vlastní lexém (atribut Frazém), případně též jeho varianta (atribut Varianta) může být opatřen speciálním klasifikátorem (atributy Stylový\_příznak a Oborový\_příznak). Taktéž grafický zápis jednotlivých atributů je shodný. Veškerá frazeologie se však neuvádí za jednotlivé významy jako kolokace, ale souborně na konci hesla po speciálním znaku rombusu ♦ (viz výše zmíněný status bezprostředního potomka pouze k elementu Lemma).

#### 2.4.9 Element Reference (defaultně *Reference*, třída *SmartReferences*)

je interním, avšak veřejným elementem, sloužícím k odkazování na jiná hesla. Je potomkem elementů Lemma, Význam, Kolokace, Frazém a Úzus, naopak jako rodič nefiguruje a nemá rovněž žádné atributy, definuje pouze různé druhy odkazů. Defaultně nastavená šablona byla podle našich potřeb předefinována, LČ se omezí na jediný typ odkazu (respektive jeho dvě varianty): a to k jinému heslu, respektive k jeho dílčímu významu, kde se uživatel o hledaném tvaru dozví více informací a jež se v hesláři nachází buď před (Viz výše ↑) hledaným heslem, anebo za ním (Viz níže ↓). Ekvivalenční reference je vyřešena vesměs atributem *Dešifrace\_zkratky* elementu Lemma, s ostatními typy odkazů (srovnej [téz]; odkazování na synonyma, antonyma či slova významově související apod.) se nepočítá, do značné míry tyto odkazy nahrazuje neveřejný atribut *Ref* elementu Lemma.

Odkazování vně hesláře na přehled gramatické gramatiky je řešeno nereferečním, textovým způsobem na úrovni jednotlivých atributů, a to pomocí komplementární šipky → a klíčového slova *Gramatika* s příslušným číslem stránky, popř. paragrafu.

#### 2.4.10 Element Úzus (nově přidáný element, třída *User*)

je veřejným elementem zaznamenávajícím poznámky o úzu (*usage notes*) a podobné prostředky především pragmatické povahy (atribut *Text*). V případě gramatické problematiky se pomocí atributu *Odkaz* odkazuje (způsobem popsaným v předešlém odstavci) vně hesláře na příslušnou pasáž gramatického přehledu na konci LČ.

#### 2.4.11 Element Excerptum (nově přidáný element, třída *User*)

je neveřejným elementem sloužícím k záznamu aktuálního úzu daného lexému. Může fungovat pouze jako potomek a lze jej vázat k elementům Lemma, Význam, Kolokace, Frazém a Ekvivalent. Zahrnuje atributy vesměs textové

povahy: Excerptum (samotný výpisek), Překlad (jeho překlad do češtiny, ať už převzatý či ad hoc stvořený), Zdroj (pramen, z něhož byl výpisek pořízen) a Poznámka (případná doplňující poznámka).

#### 2.4.12 Element Č-Lot (nově přidaný element, třída *User*)

je neveřejným potomkem elementu Lemma obsahujícím poznámky o užití slova ve specifickém česko-lotyšském kontextu.<sup>89</sup> Struktura atributů je tu podobná jako u předchozího elementu Excerptum (element Č-Lot může být chápán jako jeho zvláštní podtyp): vyplňují se atributy Text, Překlad, Zdroj a Poznámka.

#### 2.4.13 Element Ref (nově přidaný element, třída *User*)

je neveřejným potomkem elementu Lemma odkazujícím na webovou stránku s dalšími informacemi o daném hesle, často multimediální povahy.

Pro potřeby LČ nebyly použity dva elementy z defaultní šablony: element *Example* byl nahrazen elementy Kolokace, Frazém a zčásti též Excerptum a element *Definition*, typický pro jednojazyčné, výkladové slovníky, byl nahrazen elementem Ekvivalent.

Lze tedy shrnout: Většina údajů zaznamenávaných do naší databáze je veřejné povahy – měly by být tudíž zahrnuty do LČ. Kromě toho jsou tu však také informace neveřejné, sloužící interním potřebám zpracovatele. Ty lze i nadále rozšiřovat jednoduchou úpravou DTD struktury: databáze se tak potenciálně stává zdrojem nejenom LČ, ale v širším smyslu i znalostní databází o lotyštině vůbec, a jako takovou i pramenem k dalším projektům rázu lingvistického,

---

<sup>89</sup> Např. u hesla *salamandra* ‚mlok‘ zaznamenaná informace o překladu Čapkovy románu *Válka s Mloky* (*Karš ar salamandrām*), s dalšími potenciálními údaji (rok lotyšského vydání, jméno překladatele apod.).

translatologického či pedagogického, může být využita též pro účely strojového zpracování přirozeného jazyka (srov. Vossen 1992, Vondříčka 2011: iii, Čermák 1995b: 62–63 aj.).

## Část třetí

### Teoretický rámec a metodika zpracování LČ

#### 3.1 Možnosti makrostruktury

Makrostruktura slovníku bývá dnes oproti původní definici (Hausmann–Wiegand 1989: 328)<sup>90</sup> chápána širěji, vedle samotného inventáře hesel zohledňuje též jejich vzájemnou provázanost a způsob uspořádání do jednotícího celku, v obecnějším slova smyslu pak organizaci informací ve slovníku s ohledem na uživatele (Hartmann 2001: 64, Čermák 1995d: 236). V souladu s tímto pojetím a v návaznosti na předchozí kapitolu, v níž byly mimo jiné shromážděny lexikální zdroje budoucího hesláře (viz 2.3), budou v této části popsány další fáze tvorby slovníku: stanovení kritérií pro výběr hesel, sestavení hesláře a jeho systemizace (3.1.1) a rozčlenění hesel na různé typy (3.1.2).<sup>91</sup>

Někdy bývají do makrostruktury zahrnovány ještě další části slovníku, např. předmluva či přílohy (Filipec 1995: 21), ty však v této práci vydělujeme zvlášť; stručně o nich bude pojednáno v kap. 3.3.

##### 3.1.1 Heslář LČ

Prvotní verze hesláře chystaného LČ z listopadu 2008 je *corpus-driven* povahy: vznikla na základě korpusu lotyšských textů o velikosti cca 11,5 milionů textových pozic, jenž posléze posloužil jako podklad pro tvorbu LVK. Stanovením minimální frekvence 3 pro lemmatické tvary, získané morfologickou anotací, bylo vygenerováno 25 758 jednotek, jež byly úspěšně importovány do

---

<sup>90</sup> „[...] the ordered set of all lemmata (headwords)“

<sup>91</sup> Srov. „macrostructure decisions“, zahrnující „[d]eciding on the types of entry the dictionary will include, and organizing the headword list“ (Atkins – Rundell 2008: 160).

TLexu.<sup>92</sup> Daný objem odpovídal autorovým prvotním představám o velikosti slovníku: především korespondoval s rozsahem obou česko-lotyšských slovníků (viz 2.1.3). Strojově vytvořený heslář bylo pochopitelně nutno zbavit chybně vygenerovaných forem, v menší míře též hesel, u kterých bylo překročení frekvenčního prahu zapříčiněno nevyvážeností pilotního korpusu. K této revizi byl využit heslář LVV, v té době nejreprezentativnějšího výkladového slovníku lotyštiny. Zároveň byl heslář LČ doplněn o chybějící hesla, která figurovala v hesláři LVV, čímž se jeho objem přiblížil 30 tisícům položek (velikost samotného LVV). Byl tak naplněn záměr pokrýt pokud možno všechen bezpříznakový lotyšský lexikon, jenž má tvořit jádro LČ. Následným čerpáním z dalších zdrojů (viz 2.3), ale též zařazením hesel z hesláře LVV vydělených (*propria*, zkratky – viz dále 3.1.2.5) se databáze rozrostla o nových cca 5 tisíc položek.

Aktuálně čítá heslář necelých 35 tisíc hesel, přičemž se počítá s dalším nárůstem, a to o specifické vrstvy lotyšského lexikonu (zvláště o odbornou terminologii a lexémy typické pro mluvený jazyk), které do ní byly dosud zanášeny nesystematicky. Výsledný rozsah hesláře LČ by se tak oproti původním představám mohl rozrůst až na dvojnásobek a odpovídal by spíše objemu jiných aktuálních překladových slovníků s lotyštinou jako výchozím jazykem (LP 48 tisíc hesel, LL 43 tisíc).<sup>93</sup> Díky tomu by měl LČ lépe dostát výše vytyčeným cílům (viz 2.2), zvláště jde-li o druhý bod (potřeby profesionálních překladatelů a tlumočnicků pracujících s odbornými texty).

Výběr hesel do hesláře LČ se opírá primárně o kritérium kvantitativní, tj. o frekvenci lexikální jednotky, určující její pozici v jazykovém systému a od

---

<sup>92</sup> Kromě důvodů daných výše postulovanou metodologií tu nelze přehlížet ani ryze praktický aspekt, a sice značnou časovou úsporu oproti alternativnímu postupu, kdy by byl heslář převzat z některého stávajícího slovníku a musel by být zanesen do databáze manuálně.

<sup>93</sup> S rostoucím počtem slovníkových hesel úměrně vzrůstá i tzv. lexikální pokrytí textu, tedy poměr slov v textu, které lze (alespoň teoreticky) pomocí slovníku úspěšně dekodovat. F. Čermák (2010a: 245) s oporou v datech z korpusu SYN2000 udává tyto poměry: znalost 1000 nejfrekventovanějších českých slov by měla stačit k porozumění 64,9 % slov v libovolném českém textu, 10 000 slov k porozumění 91,38 % textu, procenta se dále zvyšují (20 000 – 95,53 %; 30 000 – 97,09 %; 40 000 – 97,86 %; 50 000 – 98,27 %). Pro lotyštinu jsou dostupné údaje z LVBV (13): 1000 nejfrekventovanějších slov – 64 %, 2000 – 75 %, přičemž nejvyšší pokrytí textu vykazují beletristické texty (69, resp. 78 %), nejnižší texty vědecké (55, resp. 64 %). Srov. též údaje uváděné V. Kuzinovou (Kuzina 1998a, 1998b) s maximální hodnotou 5000 slov – cca 93 % textu.

toho se i odvíjející životnost.<sup>94</sup> Nelze však vycházet pouze z četnosti, zvláště při vědomí výše naznačených slabin lotyšských korpusových dat; mohla by tak být opomenuta některá hesla, která je – i přes jejich relativní řidkost – z uživatelského hlediska důležité zařadit (Čermák 1995d: 235). Proto je kvantitativní hledisko doplněno o skupinu měřítek kvalitativních, hrajících při sestavování hesláře roli sekundární, nikoliv však podružnou. Jde o tato kritéria:

- tzv. nominativní (též pojmenovávací) reprezentativnost, tedy „výběrová úplnost pokrytí zvolení úrovně pojmenovávacích potřeb daného typu slovníku“ (ibid.: 234). Příklady uváděné F. Čermákem (pragmatické výrazy, reálie, oborová terminologie či metajazyk) doplňujeme ještě o uzavřené třídy slov, jejichž prvky musí být do hesláře zahrnuty buď všechny, anebo žádný z nich;<sup>95</sup>
- míra polysémie (Čermák 1995c: 95): s rostoucí polysémií vzrůstá i oblast lexikonu, již daný lexém pokrývá, a tudíž i potřeba zařadit jej do hesláře;
- míra derivačního potenciálu (ibid.): do hesláře je zapotřebí zahrnout též lexémy sloužící jako častý slovotvorný základ jiným lexikálním jednotkám; uživatel pak může analogicky dekodovat i jiné deriváty, ve slovníku neuvedené, čímž se rozhojňuje jeho lexikon i kompetence ve výchozím jazyce (Čermák 1995d: 235);
- invariantnost: v případě omezeného místa je nutno vybrat z vícero synonymických prostředků bezpříznakovou formu (ibid.); k označování příznakových lexémů viz dále 3.2.6;

---

<sup>94</sup> Svensén (2009: 63–64) klade na první místo autentičnost (*authenticity*) daného lexému, ta však s frekvencí do značné míry splývá – v jazyce je autentické to, co se vyskytuje běžně, opakovaně.

<sup>95</sup> Tak lze např. v ČLN z označení planet sluneční soustavy najít jen tato čtyři hesla: *Merkur* (v korpusu SYN2000, z něhož ČLN čerpal, doložen pouhými 89 lemmaty, přičemž zanedbáváme polysémii), *Mars* (1353) a *Jupiter* (649) a diskutabilně pod lemmatem *zem(ě)* (bez distinkce malého a velkého písmena) též *Země* (5837 tvarů na dotaz na slovní tvar *zem[ěi]*). Chybějící hesla mají rozdílnou frekvenci, vždy však větší než v hesláři figurující *Merkur*: *Venuše* (467), *Saturn* (295), *Uran* (121, v ČLN pouze *uran*) a (tou dobou aktuální) *Pluto* (258). V ČL nejsou kosmonyma vůbec uváděna.



- morfologická průhlednost: nepřítomnost forem nepravidelných (zejména supletivních) či nesnadno přiřaditelných k základovému slovu může uživatele splést (viz příklad v úvodu kap. 3.2.3.2);

- specifický lotyšsko-český kontext;<sup>96</sup>
- výskyt lexému v ostatních slovnících.<sup>97</sup>

Zbývá určit způsob řazení jednotlivých hesel. To bývá u dvoujazyčných slovníků téměř výhradně abecední. V LČ upouštíme od praxe hnízdování (není-li za ni považováno seskupování morfologických či pravopisných dublet téhož lexému – viz též dále 3.2.2), u víceslovných lemmat zohledňujeme mezeru, jež předchází písmenům (tzv. *word-by-word principle*); grafémy s diakritikou následují po bezdiakritických, verzálky po minuskách. Homonyma řadíme podle vzájemného poměru jejich frekvence,<sup>98</sup> jen v případě nedostatečných dat se uchylujeme k jinému principu, nejčastěji logickému (k různým možným způsobům řazení významu srov. Lew 2009, Adamska-Sałaciak 2006: 93n. aj.).

### 3.1.2 Typy hesel v LČ

Heslář LČ zahrnuje následující typy hesel:

3.1.2.1 Většinu hesláře budou tvořit standardní hesla, založená na slovních lexémech, s různou mírou komplexnosti jazykového popisu, jež by měla odrážet jejich postavení v lexikálním systému lotyštiny. Jádru lexikonu, dobře

---

<sup>96</sup> Takto by měl být do hesláře LČ zařazen historismus *prāģeris* (odvozený od propria *Prāga*), označující potulného muzikanta původem z Prahy; srov. citaci v LLVV [62/334]: *Šai laikā [17. gadsimtā] Latvijā ieceļoja daudzi vācu, čehu (tā sauktie «prāģeri»), poļu un citu tautu muzikanti*. 'Tenkrát [v 17. století] do Lotyšska přicestovalo mnoho německých, českých (takzvaní pražáci), polských a jiných muzikantů.' Ekvivalent je vhodné doplnit o patřičnou encyklopedickou glosu reflektující etymologii slova. Daný lexém je přitom zcela periferní, jako historismus jej uvádí pouze LLVV, v korpusech je doložen pouze jednou (LVK), respektive dvakrát (LVTk).

<sup>97</sup> Vzhledem k časté praxi přebírání adresářů ze starších slovníků lze toto vodítko použít pouze okrajově, často totiž bude v rozporu s objektivním frekvenčním kritériem. Platí to však i opačně, pro případ absence daného lexému ve většině slovníků: neologismy ve starých pramenech přirozeně nenajdeme, často pak ani např. slova tabuová.

<sup>98</sup> Lotyšská lexikografie upřednostňuje řazení podle slovnědruhové příslušnosti, kdy např. substantiva systematicky, nehlédě na frekvenci, předcházejí adjektivům, viz např. *potenciāls*<sup>1</sup> ('potenciál') – *potenciāls*<sup>2</sup> ('potenciální'), přestože jasně převažuje adjektivní užití (72 výskytů proti 35 v korpusu *miljons-1.0*; citováno podle Andronova – Andronovs 2011: 50). Srov. též např. v LL (267) či LP (569) řazení homonyma *ir*<sup>3</sup> (3. os. ind. prez. slovesa *būt*, 'být') až jako poslední, přestože se jedná o druhý nejfrekventovanější lotyšský tvar vůbec.

vymežitelné už na základě malých korpusů, musí být zastoupeno bezezbytku; obecně vzato je třeba frekventovaným lexémům poskytnout úměrné množství prostoru, byť by působily zdánlivě neproblematicky (zvláště internacionalismy). Standardní heslová stať zahrnuje vedle lemmatu, eventuálně jeho varianty či variant, základní charakteristiku fonologickou (výslovnost, přízvuk, intonace), gramatickou (slovnědruhovú příslušnost a morfologické tvary) a syntaktickou (valence), informaci o stylistické a/nebo oborové platnosti, dále pak část sémantickou (výčet českých ekvivalentů), exemplifikační (kolokace a frazémy s českými protějšky), případně na konci odkaz (ať už na jiné heslo či na gramatický přehled za heslářem) a/nebo informaci o úzu daného lexému (včetně korpusových statistik) – blíže k složení heslové stati viz dále kap. 3.2. Podoba hesla je pochopitelně nemálo ovlivněna jeho slovním druhem (viz např. odlišnou exemplifikaci jmen a gramatických slov, naznačenou dále v kap. 3.2.3).

Lexémy patřící k specifickým oblastem jazyka (odborná terminologie, slang aj., případně též vybrané archaismy a dialektismy) jsou zpracovány ve zkrácené formě, většinou nebývají tak bohatě exemplifikovány. Podstatné je naznačit jejich zvláštní charakter příslušným markerem.

**skelet**[s] m<sub>1</sub> 1. <anat> kostra, skelet • *člověka s. lidská kostra, kostra člověka • dzīvnieka / dinozaura s. zvīřecī / dinosauří kostra, kostra zvīřete / dinosaura • ārējais / iekšējais s. vnější / vnitřní kostra • ~a muskulī kosternī svaly* || (lidská kostra) *kostlīvec* || 2. <tech> (nosná konstrukce) kostra, skelet • ēkas s. <stav> skelet budovy || <dopr> kostra, skelet || 3. (románu, plānu ap.) kostra, osnova, náčrt || 4. <geol> • augsnes s. pūdnī skelet || ♦ s. skapī *kostlīvec ve skříni*

**skeletons** [o:] m<sub>1</sub> <sport> skeleton

**skeletonists** [o:] m<sub>1</sub> <sport> skeletonista

3.1.2.2 Zvláštní typ hesel tvoří lexémy s omezenou kolokabilitou, jež bývají často součástí frazémů.<sup>99</sup> Nemusejí být opatřeny slovnědruhovým markerem (slovnědruhovú charakteristika je buď těžko určitelná, či irelevantní) ani českým ekvivalentem (vhodnější je hledat protějšek v rovině nadslovní); znak dvojtečky

<sup>99</sup> Srov. „konstruovaná hesla“ (Čermák 1995d: 236).

následující bezprostředně po lemmatu naznačuje jeho nestandardní kombinatorku s jinými lexémy a/nebo nejednoznačný český protějšek.

**lūgtin:** ♦ l. lūgt (naléhavě, intenzivně) *prosīt, žadonīt*

**bešā:** ♦ palikt b. *zůstat s prázdnou, ostrouhat* ♦ atstāt b. (*ko*) *nechat na holičkách (koho)* ♦ viņš no tīrgus pārņāca b. *vrátil se z trhu s prázdnou*

**vakarēdiens** m<sub>1</sub> <náb>: • Svētais V. **a)** *Poslední večere [Páně] b)* *svaté přijímání*

3.1.2.3 V hesláři LČ budou významně zastoupena rovněž odkazová hesla, která zahrnují jednak tvary supletivní a nepravidelné (včetně všech hláskově alternovaných kmenů sloves 1. konjugace – viz začátek kap. 3.2.3.2), jednak dublety.<sup>100</sup> U první jmenované skupiny nemusí pouhý odkaz stačit k jednoznačné identifikaci lexému, kde je to možné, uvádíme proto i jeho přesnou morfologickou charakteristiku, jež může mnohdy ušetřit uživateli další hledání: požadovanou informaci si odvodí spojením gramatického údaje a cílového lemmatu (Svensén 2009: 131).

**iekam; iekām** konj ↓iekams

**esi** 2sg préz ↑būt

**suņa** Gsg ↑suns

Podobně vhodné je opatřit odkazová hesla, jakkoliv mají primárně referenční povahu, relevantní doplňkovou informací i v případě morfologických dublet: ty méně frekventované, často více či méně příznakové, odkazují k četnějším a/nebo neutrálním podobám; doplněny by měly být o náležitý marker, určující jejich vztah k odkazovanému lemmatu (řídce k jeho dílčímu významu), mimoto mohou být vybaveny příkladovou částí, zpravidla jen omezenou, uzuální informací apod.

Odkazové heslo může referovat k celé skupině lemmat, pomlka zastupuje zbytek lexému.

<sup>100</sup> Do jisté míry lze jako odkazová hesla chápat rovněž hesla zkratková (viz níže), která zastupují nezkrácené podoby lexémů. Mezi těmito jednotkami je vztah ekvivalence, naznačený v LČ rovnítkem (viz dále 3.1.2.5.2), uživatel však může hledat dodatečnou informaci o nezkráceném lexému u příslušného lemmatu – dešifrace zkratky tak zároveň funguje jako reference k jiné položce hesláře.

**dābol–** <<sup>?</sup>dial> ↑ābol–<sup>101</sup>

**nākoš–** ↑nākam–<sup>102</sup>

Odkazových hesel nepoužíváme v případě morfologicky nesouvisejících synonym, jako to běžně činí některé slovníky s cílem ušetřit místo – argument proti této praxi vidíme v uživatelském diskomfortu (zvláště jde-li o slovník více-svazkový, např. LR). Zabránit by se rovněž mělo dvojímu odkazu, tzn. situaci, kdy je uživatel navigován na heslo, které je jen dalším odkazovým heslem.<sup>103</sup>

Technicky je systém referencí v TLexu zajištěn pomocí interního elementu Reference (viz 2.4.9). Vedle standardních odkazů, označených v závislosti na vzájemném postavení obou hesel v hesláři značkou ↑ (viz výše) či ↓ (viz níže), lze pro potřeby autora databáze přidat další typy referencí, např. odkazování sémantické (odrážející paradigmatické vztahy mezi hesly) či slovotvorné (k označení všech lexémů spadajících do téže slovní čeledi), ty však nebudou součástí LČ.

<sup>101</sup> Odkazuje na dialektální tvary lexémů *dābols* ‚jablko‘, *dābolīš* ‚jetel‘, *dābolaine* ‚jeteliště‘, *dābolains* ‚(o koni) grošovaný‘ aj., z nichž některé jsou nicméně zahrnuty do MLVV, a to bez příslušného diatopického klasifikátoru (srov. též LP: 367, LR: I/227, LL: 148).

<sup>102</sup> Odkazuje na více než desítku hesel převážně adverbiální povahy. Při určení odkazové varianty vycházíme z frekvenčních údajů (*nākoš*. \* v LVK 120krát, v LVTK 12 950krát; *nākam*. \* v LVK 2190krát, v LVTK 30 849krát). Srov. řešení nabízené LP (856):

*nākamsvētdien* *adv.* w *nastēpnā niedzielē*  
*nākamvasar* *adv.* w *nastēpne lato, nastēpnego*  
*lata*  
*nākamziem* *adv.* w *nastēpnā zīmē, nastēpnej zimy*  
*nāk|ošais, ~ošā partic. def., patrz nākamais*<sup>1</sup>  
*nāk|ošais, ~ošā s. def., patrz nākamais*<sup>2</sup>  
*nākošdien* *adv., patrz nākamdien*  
*nākošgad* *adv., patrz nākamgad*  
*nākošmēnes* *adv., patrz nākammēnes*  
*nākošnakt* *adv., patrz nākamnakt*  
*nākošnedēļ* *adv., patrz nākamnedēļ*  
*nākošpavasār* *adv., patrz nākampavasār*  
*nākošreiz* *adv., patrz nākamreiz*  
*nākošruden* *adv., patrz nākamruden*  
*nākošsvētdien* *adv., patrz nākamsvētdien*  
*nākošvasar* *adv., patrz nākamvasar*  
*nākošziem* *adv., patrz nākamziem*  
*nākotn[e s., sg. tnt. (V) 1. przyszłość; kas gaida*

**Obrázek č. 6:** Zpracování lexémů začínajících komponentem *nākoš–* v LP

<sup>103</sup> Teoreticky si lze představit až třířázkovou referenci např. v případě zastaralého (nicméně stále ještě doloženého: v LVK 2 výskyty, v LVTK 111) tvaru *nevaids* ‚není‘, jež by uživatele navigovalo k aktuální podobě *nav*, to dále k slovesnému lemmatu *nebūt* a konečně k pozitivnímu tvaru *būt* ‚být‘. Takovýto případ je lepší řešit přímým odkazem.

**nevaids** (= nav) <zast> ↑būt

3.1.2.4 V souladu s moderními lexikografickými trendy se LČ neomezuje pouze na jednotky slovní povahy; lemmata mohou být jak víceslovná, tak mít povahu sublexému; přičemž však platí restrikce prostorové dané předpokládá-nou tištěnou formou LČ, a to zejména u víceslovných lexémů, jejichž počet mnohonásobně převyšuje množství lexémů jednoslovných.

3.1.2.4.1 V lotyšských překladových slovnících, především těch tištěných, se víceslovná hesla objevují sporadicky a jejich výběr mnohdy působí dojmem nahodilosti.<sup>104</sup> Mnohem početněji jsou zastoupena v hesláři slovníků elektronických, kde se však občas objevují lemmata až absurdní, vzniklá mechanickým převrácením výrazů ve výchozím a cílovém jazyce.<sup>105</sup> Není tu tak splněn základní předpoklad, že má-li být vícečlenné lexikální jednotce vůbec přisouzen status lemmatu, musí být tato v jazyce dostatečně ustálená, lexikalizovaná (Svensén 2009: 102–103; kritéria pro určení víceslovných lexémů podává např. Zgusta 1971: 144–152); volná spojení slov mohou posloužit jako exemplifikace, v samotném hesláři by se však vyskytovat neměla.

Pro potřeby lotyšsko-české lexikální databáze zpracováváme víceslovné lexémy různě, s ohledem na sémantickou povahu jejich komponentů. Kombinaci autosémantik (včetně proprií) zařazujeme obvykle mezi kolokace pod některou z části sousloví, případně pod obě (všechny). Spojení autosémantika se synsémantikem zpracováváme jednak jako kolokaci, ve většině případů však souběž-ně také jako samostatné lemma. Duplicita tu nevadí: databáze není, na rozdíl od budoucího slovníku, nikterak limitována a alternativní řešení se mohou snadno odfiltrovat, budou-li náležitě označena. V LČ se nicméně počítá, alespoň podle

---

<sup>104</sup> Viz např. heslo *salto mortale* v LP (1187), následující po heslu *salto*, přičemž jiná vícečlenná lemmata se v tomto slovníku vyskytují jen vzácně. Srov. též jisté nedůslednosti v ČLN: jako samostatná lemmata tu figurují např. *Jaderské moře*, jako jediné z názvů moří, (objevuje se též jako kolokace pod heslem *moře*) či *Malá Asie*, kdežto analogický výraz *Blízký / Dálný východ* už je podán jen jako kolokace (zato pod oběma příslušnými komponenty).

<sup>105</sup> Za všechny jmenujme např. heslář LE, kde lze nalézt lemma *deniņu kaula aizauss augura iekaisums* ‚zánět bradavkového výběžku spánkové kosti‘ či čtveřici hesel 1. *kakao pupa*, 2. *kakao pupas*, 3. *kakao pupiņa*, 4. *kakao pupiņas* se stejným významem (‚kakaové boby‘), odlišných pouze gramatickým číslem (singulár 1, 3 × plurál 2, 4) a přidáním rysu deminutivnosti (3, 4). (Oba příklady citovány podle Bušs 2015.)

aktuální koncepce, s prvním řešením<sup>106</sup> (srov. též dvojí nabízené řešení v kap. 4.3); víceslovná lemmata jsou v něm tak omezena pouze na kombinace dvou (či více) synsémantik, což je také nejčastější případ víceslovných hesel v lotyšských překladových slovnících. Konečně frazémy zařazujeme důsledně pod klíčový prvek na konec hesla.

Takto rozvržený postup však komplikuje aktuální ortografická konvence, přistupující odlišně k analogickým lexémům. Tak jsou např. zájmena *ikkurš* ‚každý‘ a *jebkurš* ‚kdokoliv; kterýkoliv‘, vzniklá spojením zájmena *kurš* ‚který‘ s partikulemi *ik* / *jeb*, psána dohromady, kdežto zájmena *diez kurš*, *diezin kurš*, *nez kurš*, *nezin kurš* (se společným významem ‚kdovíkdo, bůhvíkdo; kdovíkterý, bůhvíkterý‘) a *kaut kurš* ‚některý, kterýkoliv‘ se píše zvlášť, ačkoliv jsou co do složení (první část lexému je rovněž částicí) i chování stejná.<sup>107</sup> Podle výše nastíněného principu zpracováváme tato hesla oběma způsoby, přičemž v LČ se prozatím nepočítá se samostatnými hesly.

Příklady vícečlenných lemmat:

**kaut gan** konj *i když, přestože, ačkoli[v]*, *třebaže* • k. g. nebija vēls, ārā jau kļuva tumšs *i když nebylo pozdě, venku už se setmělo* • k. g. ļoti gribējās, es tomēr nekā neteicu *ačkoliv se mi moc chtělo něco říct, přesto jsem nic neřekl* • bieži esmu neizlēmīgs, k. g. no malas tā nešķiet *často bývám nerozhodný, i když zvnějšku se to nezdá*

**diezin vai; diez vai** part **1.** *sotva, těžko, stěží* • d. v. kādu pārsteigs, ka.. *stěží někoho překvapí, že...* • d. v. viņam var uzticēties *těžko se mu dá věřit* • šo plānu d. v. var saukt par racionālu *tento plán může být jen těžko nazván rozumným* • Varbūt mums kaut ko vajag? – D. v. *Nepotřebujeme něco? – Nemyslím.* || **2.** (se záporným tvarem slovesa) *pravděpodobně, asi ne* • tomēr tik pievilcīgi apstākļi d. v. saglabāsies ilglaicīgi *takto atraktivní podmínky se však asi dlouhodobě neudrží*

<sup>106</sup> Nabízí se však též zavedení dodatečného elementu Sublemma (srov. Čermák 1995d: 244), jež by bylo klonem elementu Lemma (2.4.3) a fungovalo jako jeho bezprostřední potomek. V té souvislosti by bylo praktické rozlišit též – a to i po stránce grafické – kolokace ustálené od volných (a další běžně exemplifikace). Příkladem takového sublemmatu jsou následující dvě předložková spojení, jež se co do rozsahu vyrovnají běžným jednoslovným heslům; v prvním případě nadto pro vydělení v samostatné lemma hovoří jiná slovnědruhovátá platnost daného lexému.

**jokls** m<sub>1</sub> *vtip, žert* [...] Δ **pa jokam** adv *žertem, v žertu, z legrace* • pus pa j. *napůl žertem* • ne pa j. *vážně, vědomě*

**velti** adv *marně, zbytečně, nadarmo* [...] Δ **par velti** adv **1.** *zadarmo, bezplatně* • dabūt (saņemt) (*ko*) p. v. *dostat (co) zadarmo* • p. v. pieejams *bezplatně přístupný (dostupný)* || **2.** = *velti* • tu p. v. pūlējies *snažil (namáhal) ses zbytečně* • ne p. v. sacīts, ka.. *ne nadarmo se říká, že...*

<sup>107</sup> Srov. též kombinaci dvou zájmen *kurš katrs* ‚kdekdo; kdekterý‘, eventuálně i jiný případ, kdy se kombinují dvě částice: *diez vai* ‚sotva, těžko‘ × *diezgan* ‚skoro, téměř‘.

**ak vai** interj (vyjadřuje zklamání, nemilé překvapení, úlek apod. emoce) *ach jo, a jéjé, jejda, šman-kote*

3.1.2.4.2 Sublexikální lemmata budou v hesláři LČ reprezentována jednak uzavřenou množinou lotyšských prefixů, jednak produktivními prvními částmi složenin – ty jsou vybaveny i pár příklady, jež mají tuto produktivnost naznačit a zároveň pomoci uživateli správně porozumět okazionalně vytvořeným kompozitům, které do hesláře zahrnuty nejsou. Vzhledem ke skutečnosti, že se uživatel při dekódování orientuje primárně podle začátku, nikoliv konce lexému, nezařazujeme do hesláře jiné afixy (sufixy, koncovky),<sup>108</sup> jež mohou být ve stručnosti probrány v gramatickém přehledu.

**jā–** prefix debitivu, ve spojení s 3. osobou přítomného času vyjadřuje nutnost, nezbytnost, potřebu ap., např. *man [ir] jāiet* („musím jít“), *tev ir mani jāapciemo* („musíš mě navštívit“), *viņai būs jā atrod citu vietu* („bude muset najít jiné místo“) → **Gramatika: s. 1052**<sup>109</sup>

3.1.2.5 V lotyšských překladových slovnících se dosud udržuje praxe vyčleňovat z hlavního hesláře specifické skupiny hesel (zkratky, zeměpisná a osobní jména, případně též jiné); ty tvoří samostatné dílčí hesláře, jež bývají součástí příloh. To je koncepce problematická (a dnes už dosti překonaná), jednak kvůli zjevné nesystémovosti, jednak též z hlediska uživatele.<sup>110</sup> V LČ od ní proto upouštíme, ve prospěch jediného hesláře, zastřešujícího veškerou makrostrukturu. Takto v něm budou zahrnuty následující třídy slov:

3.1.2.5.1 *Propria*, konkrétně: antroponyma (soudpis nejfrekventovanějších lotyšských křestních mužských i ženských jmen, ne však už příjmení), doplněná

<sup>108</sup> Opět nám však nic nebrání zaznamenávat do TLexu i tyto formy; jednoduchým filtrováním je pak lze eliminovat z hesláře LČ, přičemž v databázi zůstanou. Příklad takového hesla, zobecňujícího zvrtná feminina (3.2.3.1.4.7)

**–anlās** *f<sub>(refl)</sub>* (~ās, A/I sg, Gpl ~os) zakončení reflexivních feminin, např. *vēlēsānās* („přání, touha“), *vingrināšanās* („cvičení, trénování, trénink“), *precēsānās* („ženitba, vdavky“).

<sup>109</sup> Jde o odkaz na gramatický přehled za heslářem, kde je daný jev podrobněji vyložen. Číslo stránky je smyšlené. Viz též dále 3.2.8.

<sup>110</sup> Příkře se k této praxi vyjadřují Atkins – Rundell 2008: 179 („[...] from the point of view of users [...] it was simply another obscure idiosyncrasy of dictionary editors.“). Dlužno poznamenat, že lotyšská lexikografie zůstává tomuto postupu věrna (oddělené hesláře zeměpisných jmen v LP, LL i LA; úplná absence *proprii* v LR aj.); v tomto ohledu je zapotřebí ocenit rozhodnutí S. Nikulcevojové v ČLN tuto praxi, navzdory tradici, opustit.

o vybraná zoonyma; dále toponyma (včetně kosmonym) a některá chrémantonyma. Při výběru těchto lexémů do hesláře se opíráme o reprezentativní rejstříky proprií, hlavním kritériem je vedle frekvence také nominativní reprezentativnost (viz výše 3.1.1), ale též mimojazykové skutečnosti (např. velikost, respektive geografická poloha měst, jejich relevance pro český či baltický kontext apod.), jež se oproti běžným heslům projevují zvýšenou mírou encyklopedických informací a opisných ekvivalentů. Od apelativ jsou propria snadno rozpoznatelná velkým písmenem, ve snaze nerozptylovat zbytečně uživatele koncentrací tudíž nevyužíváme možnost označovat je explicitně vhodnou zkratkou či značkou; jinak jsou však opatřena standardním souborem informací.

Zpracováváme především nejfrekventovanější lotyšská křestní jména, dále pak vybraná jména postav mytologických (včetně biblických), pohádkových aj.

**Jēkab|s** [æ:] m<sub>1</sub> (Vsg ~!) lot. mužské křestní jméno (≡ č. Jakub)

**Zan|e** f<sub>5</sub> (Vsg ~e!) lot. ženské křestní jméno

**Bud|ja** m<sub>4</sub> (Dsg ~am) <mytol> *Buddha* • ~as statuža / tēls *soška* / *obraz Buddha*

**Ījab|s** m<sub>1</sub> <bibl> *Job* • ~a grāmata *Kniha Jobova* || ♦ ~a vēsts *Jobova zvēst (zpráva)*

**Ahille|js** m<sub>1</sub> <mytol> *Achilleus* • ~a cīpsla <anat> *Achillova šlacha* (tendo Achilles), *achilovka* || ♦ ~a papēdis *Achillova pata*

**Neptūn|s** m<sub>1</sub> **1.** <astr> *Neptun* • ~a pavadoņi *mēsiņe Neptunu* || **2.** <mytol> *Neptun* (římský bůh moře)

**Īkšķī|tis** m<sub>2</sub> (~ša) (pohádková postava) *Paleček*

Ze zeměpisných jmen do hesláře zahrnujeme ve slovnících a zeměpisných atlasech běžně uváděné lexémy včetně víceslovných, s případnou upřesňující glosou a/nebo jinojazyčným ekvivalentem:

**Kanāriju salas** f<sub>4</sub> pl *Kanárské ostrovy*

**Adrijas jūra** f<sub>4</sub> *Jaderské moře, Jadran*

**Mičigana** f<sub>4</sub> (stát USA) *Michigan*

**Mičigans** m<sub>1</sub> *Michiganské jezero*

**Abesīnija** f<sub>4</sub> <hist> *Habeš* (dnešní Etiopie)

**Kēlne** f<sub>5</sub> *Kolín nad Rýnem* (Köln)



Lotyšská zeměpisná hesla jsou spíše encyklopedického rázu: uvádíme jejich základní charakteristiku a lokaci, eventuálně též český historický název, je-li v češtině zaveden.<sup>111</sup>

**Ventspils** f<sub>6</sub> přístavní město v sz. Lotyšsku; správní středisko stejnojmenného kraje || <hist> *Vindava*  
• ~s Augskola vysoká škola ve Ventspilsu

**Abava** f<sub>4</sub> řeka v záp. Lotyšsku, pravý přítok Venty

**Gaiziņkalns** m<sub>1</sub>; **Gaiziņš** m<sub>1</sub> kopec ve stř. Lotyšsku, nejvyšší bod LR (311,5 m. n. m.)

Obdobně jsou zpracována další propria spadající do lotyšských reálií:

**Lido** [o] m<sub>0</sub> řetězec lotyšských samoobslužných restaurací • aiziet / aizbraukt uz L. *zajít / zajet do Lida (Lido)* • pašēt L. *najíst se v Lidu (Lido)*

**Milda** f<sub>4</sub> <hov> Pomník svobody v Rize; socha ženy na tomto pomníku

Nepříliš četné případy polysémie či homonymie a apelativizace (případně též proprializace) jsou zpracovávány následně:

**Elba**<sup>1</sup> f<sub>4</sub> *Labe*

**Elba**<sup>2</sup> f<sub>4</sub> *Elba* (ostrov ve Středozemním moři)

**Mekļa** f<sub>4</sub> **1.** *Mekka* • svētceļojums uz ~u *pout' do Mekky* || **2.** m. *mekka* • modes m. *mekka módy, módní mekka*

**sarma** f<sub>4</sub> **1.** *jinovatka, námraza* || **2.** S. lot. ženské křestní jméno

Další ukázky propriálních hesel a potenciálního popisu jejich kolokability:

**Hanzļa** f<sub>4</sub> <hist> *hanza* (ve 13.–17. stol. obchodní spolek severoněmeckých a pobaltských měst) • ~as savienība *hanzovní spolek* • ~as pilsēta *hanzovní město*

**Helovīņļs** [o] m<sub>1</sub> *Halloween* • ~a vakars *halloweenský večer* • ~a maska *halloweenská maska*

**Titāņikļs** m<sub>1</sub> *Titanik* • ~a vraks *vrak Titaniku*

**Vikipēdijļa** f<sub>4</sub> *Wikipedie* • ~as raksts *článek na Wikipedii*

<sup>111</sup> Srov. zkušenost s exonymy z připravovaného Velkého německo-českého akademického slovníku (Greule 2008) či nápadité řešení v LP (1500).

Na rozdíl od češtiny nejsou v lotyštině za propria považována obyvatelská jména, píšící se malým písmenem:

**lietuvie|tis** m<sub>2</sub> (~ša) *Litevec* [...]

**parīzie|te** f<sub>5</sub> (Gpl ~šu) *Pařížanka*

3.1.2.5.2 Abreviatury s výjimkou mezinárodních (značky chemických prvků, fyzikálních veličin apod.), přebírané ze specializovaných příruček (zejména Bankavs 2003 s více než tisícovkou zkratk) a heslářů jiných lotyšských slovníků a ověřované v korpusech. Zkratkové heslo má v LČ specifický způsob zápisu: bezprostředně za lemmatem následuje v závorce nezkrácená podoba<sup>112</sup> (atribut *Dešifrce\_zkratky* – 2.4.3.12) a její český protějšek, eventuálně též s příslušnou českou zkratkou, je-li dostatečně ustálena. Morfologickou značku pro její zjevnost neuvádíme.

**piem.** (= piemēram) *například (např.)*

**b-ka** (= bibliotēka) *knihovna*

**k/t** (= kinoteātris) *kino*

**gs.** (= gadsimts) *století (stol.)* • 20. gs. beigās *na konci 20. stol.* • kopš XV gs. *od 15. stol.*  
• līdz 19./20. gs. mijai *do přelomu 19. a 20. století*

U iniciálových zkratk je zapotřebí uvádět způsob, jakým se zkratka vyslovuje – zda po jednotlivých písmenech, nebo akronymně, zvláště pak je-li situace v lotyštině jiná než v češtině (v takovém případě jsou do hesláře LČ zařazeny i zkratky cizího původu).

**ASV** [a: es ve:] (= Amerikas Savienotās Valstis) *Spojené státy americké (USA)* • ASV dolārs *americký dolar* • ASV valsts *americký stát, stát USA* • ASV vēstniecība *americké velvyslanectví, velvyslanectví (ambasáda) USA*

**HIV** [chiu] (= cilvēka imūndeficīta vīruss; human immunodeficiency virus) *virus lidské imunitní nedostatečnosti (HIV)* • HIV tests *HIV test* • HIV pozitīvs *HIV pozitivní* • HIV inficētais *[člověk] infikovaný HIV*

<sup>112</sup> Přesněji řečeno, reprezentativní tvar zastupovaného lexému. Takto zastupuje v příkladu uváděná zkratka *gs.* ovšemže celé paradigma lexému *gadsimts*, což se dá dobře naznačit vhodnou exemplifikací, kde budou figurovat i jiné pádové formy než *Nsg.*

Při zpracování zkratkových hesel narážíme na stejné problematické jevy jako u hesel standardních, jako jsou homonymie i polysémie, variantnost,<sup>113</sup> vícečlenná lemmata,<sup>114</sup> specifická valence aj. Ukazuje se, že lexikografická řešení přejatá z běžných hesel jsou dostatečně uspokojivá, malé zastoupení v korpusech však bohužel nepodává postačující data k frekvenčně danému řazení jednotlivých významů.<sup>115</sup>

**p. k.**<sup>1</sup> [pe: ka:] (= *pēc kārtas*) **1.** *po sobě, za sebou* || **2.** *střídavě*

**p. k.**<sup>2</sup> [pe: ka:] (= *personas kods*) *osobní kód* (odpovídá č. rodnému číslu)

**p. k.**<sup>3</sup>; **p/k** [pe: ka:] (= *pasta kastīte*) *poštovní schránka (příhrádka), P. O. Box*

**red. 1.** (= *rediģēj[is, ~usi]*) *redigoval[a], zredigoval[a]* || **2.** (= redakcija) *redakce* || **3.** (= *rediģēts*) *zredigováno*

Od zkratek lišíme slova zkratková a zkrácená, v případě dostatečné míry výskytu či relevance je zahrnujeme do hesláře jako standardní hesla.

**Gulag[is]** <sov> **1.** *Gulag* (Hlavní správa nápravně-pracovních táborů; součást sovětské tajné policie spravující systém pracovních táborů v SSSR) • ~a *nomettes / upuri tábory / obēti Gulagu* • ~u *pārdzīvojušie ti, kteří přežili Gulag* • *izsūtīt (ko) uz ~u poslat (koho) do Gulagu* || **2.** *g. gulag* (sít těchto táborů; jednotlivý tábor), *lāgr* • *viņiem izdevies idzīvot Sibīrijas ~os un atgriezties Latvijā*  *podařilo se jim přežít v sibiřských lágrech a vrátit se do Lotyšska*

### 3.2 Možnosti mikrostruktury

Pojem mikrostruktury je, podobně jako termín makrostruktura, chápán různými lexikografy různě. V této práci vycházíme z tradičního pojetí J. Filipce, jenž mikrostrukturu definuje takto:

<sup>113</sup> Ta je dána nejen různou měrou krácení výchozího lexému (pak by šlo však, kvůli odlišnému abecednímu řazení, o různá lemmata – např. *pers. k.* a *p. k.*, zastupující sousloví *personas kods* „osobní kód“), ale i rozličnými způsoby grafického zápisu: verzálkami, či minuskami, pomocí různých interpunkčních znamének (*a. k.*, *p/a*, *b-ba*).

<sup>114</sup> Ta píšeme v souladu s reprezentativními příručkami (Blinkena 2009, Bankavs 2003, LVPPV) s mezerou po tečce, heslo pak řadíme podle tzv. *word-by-word* principu, kdy má mezera přednost před grafémy. Příkladová zkratková hesla *p. k.* tak budou v hesláři hned na začátku písmene P, nikoliv až mezi lexémy začínajícími na *pj-* a *pl-*.

<sup>115</sup> Lze se nouzově opřít o četnost zastupovaných lexémů a finální řazení konzultovat s rodilým mluvčím. Výskyty v LVK: *pēc kārtas* 179, *personas kods* 107, *pasta kastīte* 6.

„Mikrostruktura je lingvisticky relevantní soubor informací o heslovém slově (vstupu) a jeho manifestacích. Je tedy jednotou *langue* a *parole*, jednotou paradigmaticko-syntagmatickou. Tato jednota je dynamická a spojitá, ovšem její zobrazení je nutně diskrétní. V mikrostruktuře se sdružují a prostupují roviny fonologická, grafická, gramatická (morfologická a syntaktická), sémantická, pragmatická, stylistická a textová, popř. i kvantitativní (statistická, frekvenční), historická (etymologický údaj) a rovina mezilexémových vztahů.“ (Filipec 1995: 23)

Na lexém (respektive heslo, kterým je ve slovníku reprezentováno) nahlížíme jako na jazykový znak, mající svou stránku výrazovou (tj. určitou formu) a obsahovou (tj. jistý význam). Od nich pak lze odvodit dvojí základní charakteristiku lexému: formální (zahrnující grafickou podobu, výslovnost a morfologii) a sémantickou (význam lexému, v případě dvoujazyčného slovníku prostřednictvím ekvivalentu v cílovém jazyce). Nadto lze lexém charakterizovat i z dalších hledisek: např. funkčního (způsob, jímž se kombinuje s jinými lexémy), tedy s ohledem na jeho syntagmatiku a úlohu ve vyšší struktuře, pragmatického (vztah lexémů k mluvčím, resp. interpretům), frekvenčního (četnost lexému, určující jeho postavení v lexikonu), ve slovnících v menší míře též encyklopedického (orientace na denotát, tedy vně jazyka); v závislosti na typu a cílech slovníku případně též jiných.

Obsah standardní heslové stati byl nastíněn již výše (viz 3.1.2.1). Níže předkládaná klasifikace typů informací, jimiž je lotyšský lexikon v LČ popsán, vychází do značné míry z jednotlivých jazykových rovin a také odráží strukturu hesla, jak byla prve nadefinována v TLexu (viz výše 2.4; srov. koncepci vstupních dat jakožto podkladu k mikrostruktuře – Vachková 2011: 35–36).

### 3.2.1 Lemma

Lemma (viz 2.4.3) představuje reprezentativní grafickou podobu lexému a z hlediska výstavby slovníku jde o klíčový prvek: jednak je spojnicí jeho dvou základních os (vertikální, představující heslář a potažmo i *langue*, a horizontální,

zastupovanou heslovými statěmi a odpovídající *parole*), jednak se jeho prostřednictvím ve slovníku orientuje uživatel; k hledané informaci dospívá právě skrze lemma. Není tudíž dobré tuto vstupní složku hesla informačně přetěžovat, jak se však nezřídka děje. V LČ je uvedena vlastní podoba lemmatu; jediným metajazykovým prvkem je svislice, označující místo, od kterého se slovo ohýbá (a to jen v případě, že k takovému ohýbání v rámci daného hesla skutečně dochází)<sup>116</sup> a index označující homonymní hesla (viz 3.1.1).

Lemma může být doplněno o pravopisnou či morfologickou variantu (varianty), pokud ovšem není narušeno abecední řazení hesel (v opačném případě je použito odkazové heslo – oba příklady viz 4.2). Ideální vhled do povahy variantních forem a jejich vztahu k bezpříznakovému invariantu poskytují korpusy, zejména po stránce frekvenční (poměry vzájemných četností),<sup>117</sup> ale též syntagmatické (potenciálně odlišné kolokáty). Každý symptom specifického užití by měl být podchycen příslušným markerem, v případě silné příznakovosti či malého doložení v korpusech je třeba uvažovat o úplném vyřazení ze slovníku, byť je dotyčná forma zachycena jinými slovníky – jinak by tím byly při životě uměle drženy lexémy neužívané a v důsledku potlačován i přirozený vývoj jazykového úzu a od něj se odvíjející normy.

### 3.2.2 Informace fonologická

Informace o zvukové stránce lexémů nebývá v lotyšských překladových slovnících zpravidla uváděna, v ojedinělých případech (LR, LL) jen u homograf.<sup>118</sup> Lotyšský pravopis je sice založen především na morfonologickém prin-

---

<sup>116</sup> Podobně jako v jiných tištěných slovnících jde i zde o prostředek sloužící k úspoře místa – svislice je v LČ chápána jako značka ryze lexikografická, byť leckdy může zároveň naznačit morfematickou strukturu lexému. V elektronické verzi slovníku by se od zkrácené formy zápisu flektivních forem mohlo úplně upustit, vypisovaly by se v celku – tedy tak, jak se s nimi uživatel přirozeně setká v textu.

<sup>117</sup> Takto zkoumají Andronova – Andronov (2011: 46) v lotyšských korpusech zastoupení dublet *islams* × *islāms* (spolu s odvozeninami), *doktorands* × *doktorants*, *albums* × *albūms* a *balzams* × *balzāms*. U první dvojice ustupuje původní varianta s krátkým *a*, doložená v LLVV i LVPPV, formě *islāms* (v korpusu *miljons-1.0* poměr 1:9, v LVK 10:100, v LVTK jsou data kvůli četným nelotyšským textům neprůkazná), která je – jako jediná – uvedena v novějších výkladových slovnících LVV i MLVV, převažuje též ve slovnících překladových.

<sup>118</sup> Nejexplicitnější informaci v tomto ohledu nabízí materiál Andronovův (Andronov 2002; srov. též Andronovs 1998: 141–142), uvádějící u všech hesel důsledně nejen výslovnost, ale též intonaci. Toto pojetí je však pro nás

cipu<sup>119</sup> a nepříliš četné výjimky se dají relativně snadno a systematicky popsat mimo lemma, odkázat uživatele plně na takovýto výčet vně hesláře však není optimální. Existují případy (jsou popsány dále), kdy by fonologické údaje bylo vhodné vtělit přímo do mikrostruktury jednotlivých hesel. Zároveň by však výslovnostní informace neměla být nadměrná, LČ je koncipován primárně jako překladový slovník (nemá tedy suplovat mimo jiné slovník ortoepický). Popis fonologického systému lotyštiny,<sup>120</sup> jenž bude součástí závěrečného přehledu lotyšské gramatiky, je pouhým fundamentem; v mikrostruktuře slovníku budou zohledněny vedle dublet jednak jevy rozlišující různé významy, jednak potenciálně problematické pro českého uživatele, které však nespádají do obecně popsaných zásad. Východiskem pro takovýto popis i pro jeho aplikaci v hesláři LČ je pro nás aktuální celouzemní ortoepická norma lotyštiny, reprezentovaná ortoepickým a ortografickým slovníkem (LVPPV), po stránce teoretické nejnovější lotyšská gramatika (LVG) spolu se specializovanými monografiemi (Laua 1997 aj.). Upouštíme od dialektálních odchylek, jež se mohou projevovat v idiolektech jednotlivých mluvčích, a také od alofonických realizací některých fonémů (zadopatrové /n/, palatalizované realizace fonémů /r/<sup>121</sup> a /ch/, glotální [h] jako varianta /ch/<sup>122</sup>). Korpusová data bohužel chybějí, mluvené korpusy lotyštiny neexistují (Auziņa 2010); případné dublety a odchylky od normy tak přebíráme od výše jmenovaných zdrojů, konzultujeme je s rodilými mluvčími.

---

nepřijatelné, neboť podřizuje abecední řazení hesláře fonologické transkripci, navíc nevhodně zaznamenává fonologické údaje přímo do lemmatu (srov. též podobnou praxi v ČLit, kde není jasné, zvláště nepokročilým uživatelům, která diakritická znaménka jsou součástí standardní litevské grafémiky a která označují jen slabičnou intonaci a v běžných textech se neobjevují).

<sup>119</sup> Okrajově, v mnohem menší míře než v češtině se uplatňuje též princip tradiční: transkribují a transliterují se i cizí propria, což může českému uživateli ztěžovat jejich identifikaci (Laua 1997: 143–145, 147–18).

<sup>120</sup> Půjde nutně o pouhý výtah, limitovaný nejen prostorem, ale venkoncem i předpokládaným uživatelem. Příhodný vzor co do rozsahu i formy nabízí ve svém slovníku S. Nikulceová (ČLN: 383–385). Takový výklad je ovšemže ideální založit na kontrastivním přístupu, tj. na porovnání s češtinou: paralelní jevy existující v obou jazycích (stálý přízvuk na první slabice, zánik měkkostní korelace, znělostní a artikulační asimilace konsonantů aj.) mohou českému uživateli pomoci lépe pochopit fonologický systém lotyštiny.

<sup>121</sup> Grafém *r* byl v Lotyšsku zrušen v rámci pravopisné reformy z r. 1946, nadále jej však používají Lotyši žijící v exilu (Kuškis et al. 1998: 127, 2004; LVPPV: 35).

<sup>122</sup> Foném /ch/ je od roku 1957 zapisován grafémem *h*, v exilové variantě lotyštiny se zachoval původní spřežkový zápis (*ch*) (Kuškis et al. 1998: 127).

Údaje o výslovnosti jsou ve zkrácené podobě zaznamenávány pomocí jednoduché fonologické transkripce do hranatých závorek,<sup>123</sup> přizpůsobené českému uživateli,<sup>124</sup> do samostatného atributu Fonologie elementu Lemma, případně Varianta (viz 2.4.4), objevit se mohou též v rámci atributu Morfologie (viz níže příklad *nest*); nejsou tedy nevhodně směřovány s lemmatem samotným, a obě roviny – grafémická a fonologická – jsou dostatečně odděleny, jak je patrné na četných příkladech níže.

V mikrostruktuře LČ jsou zohledněny následující případy (s výjimkou bodu 4 též s možnými dubletami), srov. též Andronovs 1998: 141–142.

3.2.2.1 Dvojí možná realizace grafémů *e/ē* (úzký a široký vokál: /e/, /e:/ × /æ/, /æ:/), přičemž úzká varianta není obvykle označena, jen ve zvláštních případech umožňujících nežádoucí interpretaci (více o těchto grafémů v lemmatu nebo výskyt homograf).

**zeme** f<sub>5</sub> (Gpl ~mju) **1. země, půda, zemina** [...]

**těvzeme** [æ:-e-e] f<sub>5</sub> (Gpl ~mju) **vlast, otčina** [...]

**nest** v<sub>1</sub> (nesu [æ], nes, nes [æ]; nesu; fut nesišu) **1. nést, nosit** [...]

**Cē|sis** [e:] f<sub>6</sub> pl (Gpl ~su [æ:/e:]) město v sev. Lotyšsku; správní středisko stejnojmenného kraje

**Cesvaine** [æ-e] f<sub>5</sub> město ve vých. Lotyšsku; správní středisko stejnojmenného kraje

**esmu** [æ] 1sg prez ↑būt

**resnvēderis** m<sub>2</sub> [æ-e:-e] **tlust'och, tloušťík, pupkáč**

**gredzenveida** [æ-æ-ei] G<sub>sg</sub>; **gredzenveidīgs** [æ-æ-ei] adj **kruhovitý, prstencovitý**

**mēr|s**<sup>1</sup> [æ:] m<sub>1</sub> **1. míra** [...]

**mēr|s**<sup>2</sup> [e:] m<sub>1</sub> **starosta, (velkého, zvl. hlavního města též) primátor** [...]

<sup>123</sup> Tento protimluv si uvědomujeme, je však dán ohledem na laického uživatele, pro nějž jsou hranaté závorky nápadnější a známější signál výslovnostní informace než dvojice lomítek.

<sup>124</sup> Mezinárodní fonetickou abecedu IPA nepoužíváme, neboť zavádí znaky, jež by uživatel (nelingvista, tím méně fonetik) mohl mylně interpretovat, např. [c] pro /t'/ (jemuž odpovídá lotyšský grafém *k*) nebo [x] pro /ch/ (zapisované v lotyštině grafémem *h*); srov. též LVG: 24. Z abecedy IPA přebíráme pouze znak pro široké /e/ [æ] (v lotyšské tradici označované písmenem *ē*, které však může evokovat polskou nosovku) a hláskovou délku označovanou dvojtečkou.

3.2.2.2 Trojí možná realizace grafému *o*<sup>125</sup> (krátký, či dlouhý monoftong, diftong: /o/ × /o:/ × /uo/), přičemž není obvykle označována diftongická varianta, pokud si to opět nevyžaduje přehlednost.

**ozol|s** m<sub>1</sub> <bot> *dub* (Quercus) [...]

**foto** [o-o:] m<sub>0</sub> <hov> *fotka* [...]

**mikrorajon|s** [o-o:/o-uo] m<sub>1</sub> **1.** *sídliště* [...]

**motorvagon|s** [o-o:-uo] m<sub>1</sub> <žel> *motorový vůz (vagon)*

**klon|s**<sup>1</sup> [o] m<sub>1</sub> **1.** <bio|> *klon* [...]

**klon|s**<sup>2</sup> [uo] m<sub>1</sub> **1.** *podlaha* (obv. z udusané hlíny či vydlážděná) **2.** *mlat, humno* [...]

3.2.2.3 Přízvuk nacházející se na jiné než první slabice lexému, nejde-li však o příznakovou, emfatickou výslovnost (příklady viz LVG: 124; Laua 1997: 109–110). Přízvučnou slabiku uvozujeme znakem apostrofu '.

**labvakar** [-'va-] interj (pozdrav) *dobrý večer*

**pamazītēm** ['pa-/-'zi:-]; **pamazītiņām** [-'zi:-] adv *pomaličku, pomalinku*

**diemžēl** ['diemžæ:l/diem'žæ:l]] part *bohužel* [...]

**Borneo** [bor'neo:] f<sub>0</sub> *Borneo, Kalimantan*

V ojedinělých případech může přízvuk na odlišných slabikách dokonce lišit různé významy, ty jsou zpracovány pod samostatnými hesly jako homografa.

**patiesi**<sup>1</sup> [-'tie-] part *skutečně, opravdu* [...]

**patiesi**<sup>2</sup> ['pa-] adv *opravdově* [...]

Vedlejší (pomocný: *palīguzsvars* – LVG: 121–122) přízvuk nezaznamenáváme.

3.2.2.4 Heterosylabická povaha spojení dvou samohlásek, jež mohou v lotyštině fungovat jako diftong,<sup>126</sup> nebo oddělená realizace kombinace souhlásek *ds*, *ts*, *dz*, *dž*. Hranici slabik označujeme znakem zvýšené tečky ·.

<sup>125</sup> Do pravopisné reformy z r. 1946 byl součástí lotyšské abecedy též grafém *Ō ō*, jenž se v exilové varietě lotyštiny udržel dodnes (Kuškis et al. 1998: 127).



**orientier|is** [ori·e-] m<sub>2</sub> *orientační bod* [...]

**saindē|ties** [a·i] v<sub>2</sub> (~jos, ~jies, ~jas; ~jos) (*ar ko*) *otrāvit se (čím)* [...]

**slidzābak|i** [d·z] m<sub>1</sub> pl *lyžařské boty, lyžáky* [...]

**pretsitien|s** [t·s] m<sub>1</sub> *protiúder* [...]

3.2.2.5 Slabičná intonace v případě povinné realizace (vesměs u homograf). Ve spisovné lotyštině (respektive středolotyšském dialektu, z nějž spisovný jazyk vznikl) existuje původní trojintonační systém, rozlišující intonaci taženou (*stieptā intonācija*), klesavou (*krītošā intonācija*) a lomenou (*lauztā intonācija*);<sup>127</sup> jejich použití je však vzhledem k nestabilní normě fakultativní, až na nečetné a snadno podchytitelné případy (Grīse 2008), kdy je na intonaci závislý význam lexému (Markus – Auziņa 2013: 20; Laua 1997: 107). V LČ budou jednotlivé slabičné intonace rozlišovány pomocí zavedených grafických symbolů: tildou ~ pro taženou, gravisem ` pro klesavou a cirkumflexem ^ pro lomenou intonaci.

**tā<sup>1</sup>** [ā] adv *tak* [...]

**tā<sup>2</sup>** [ā] pron *ta, ona* [...]

**tā<sup>3</sup>** [ā] Gsg ↓ *tas*

3.2.2.6 Nadto v omezené míře zaznamenáváme také hláskové změny, jež jsou v lotyštině dvojí povahy. Změny historické (paradigmatické), zachycené písmem jako morfonologické alternace (LVG: 89–95; Laua 1997: 69–71), jsou v LČ logicky uvedeny bezesbytku: blíže je charakterizujeme v kapitole o gramatické informaci (viz 3.2.1 a 3.2.2); zčásti fungují též jako slovotvorný prostředek – deriváty pak uvádíme jako samostatná hesla. Změny poziční, mající charakter

<sup>126</sup> Lotyština oproti češtině disponuje mnohem početnějším rejstříkem dvouhlásek: *au, ai, ei, eu, iu, ui, oi, ou, ie, uo* (psána *o*).

<sup>127</sup> LVG: 104, Laua 1997: 104–105; srov. Erhart 1984: 27 se synonymickými termíny: tažená / stoupavá a přerývaná / lomená.

Tradičně je lotyšské území dialektology děleno na sedm intonačních skupin se čtyřmi druhy slabičných intonací (čtvrtá – stoupavá [*kāpjošā intonācija*] – je vzácná, omezuje se jen na část východolotyšského nářečí [Laua 1997: 103–104]), přičemž každá skupina má svůj vlastní intonační systém, s výjimkou středolotyšského dialektu dvoučlenný, vzniklý splynutím dvou odlišných intonací. Nivelizační tendence je však patrná i ve spisovném jazyce: dochází, mimo jiné v důsledku značně rozvolněné normy (např. ve srovnání s litevštinou), k postupnému vytlačování klesavé intonace lomenou; proti tažené intonaci tak stojí netažená (ibid.: 105, 107; LVG: 105; LVP-PV: 24–25; Erhart 1984: 27, ČLN: 385).

čistě fonologický a obecně písmem nereflektované (LVG: 80–89; Laua 1997: 68–69), budou popsány v gramatickém přehledu, v samotném hesláři je však neuvádíme, s výjimkou výše nastíněné úzké, či široké realizace vokálu /e/ a několika málo zvláštních případů, omezujících se vždy na nepočetnou skupinu lexémů (samohlásková kontrakce, redukce souhláskových skupin ve starých složeninách aj.).

**neesmu** [æ:] 1sg prez ↑būt

**kooperācij|a** [o·o/o:] f<sub>4</sub> *kooperace, spolupráce, součinnost* [...]

**svētdien|a** [e:d] f<sub>4</sub> *neděle* [...]

### 3.2.3 Informace gramatická

Gramatická informace o lexému v LČ je dvojího typu, zaznamenávaná v TLexu do dvou atributů elementu Lemma, případně Význam: a sice Slovní\_druh a Morfologie (viz 2.4.3.5–6 a 2.4.5.3–4). První atribut má funkci kategorizační – přiřazuje dané lemma (význam) k určité třídě slov, zejména k jednotlivým slovním druhům, případně specifikuje některou z gramatických kategorií (prezens sloves, nesklonná jména apod.); druhý atribut uvádí vybrané tvary lexému, ať už klíčové pro dané paradigma, nepravidelné či (potenciálně) problematické z hlediska českého uživatele LČ. Kombinací těchto dvou atributů by uživatel měl jednoznačně zrekonstruovat kompletní flektivní paradigma u ohebných slov (viz též souhrn jednotlivých paradigmat v gramatickém přehledu za heslářem). Tvary, jež by striktně vzato mohly být chápány jako redundantní (např. čtyři tvary u všech osobních sloves, a to i u „neproblematické“ 2. konjugace – viz dále 3.2.3.2.4.2), uvádíme v rámci jednotnosti popisu a také jako určitý vstřícný krok vůči uživateli, jakkoliv je vyšší míra explicitnosti podmíněna většími nároky na místo.<sup>128</sup>

<sup>128</sup> Elektronické a online slovníky, prostorově nelimitované, mnohdy uvádějí kompletní paradigmatata ohebných lemmat, nejčastěji automaticky generovaná. V tištěných slovnících je nezbytné paradigma redukovat na několik

Naopak, s ohledem na uživatele lotyšského budou vybrané české ekvivalenty morfologicky vymezeny atributem Morfologie elementu Ekvivalent (2.4.6.6): takto budou substantiva charakterizována z hlediska rodu (nadto se rozlišuje ne/životnost maskulin) a čísla (singularia / pluralia tantum), verba pak po stránce vidové. Upouštíme však, vzhledem k omezenému prostoru, od uvádění morfologických forem českých ekvivalentů.<sup>129</sup> Na případnou odlišnou slovnědruhovou platnost lotyšského lexému a jeho českého ekvivalentu by měl být uživatel explicitně upozorněn.

Při morfologické klasifikaci lotyšského lexikonu se opíráme o slovnědruhovou platnost lexémů. Ta se v lotyšské jazykovědě určuje na základě čtyř kritérií: sémantického (obecně chápaný věcný význam lexému), slovotvorného (afixy charakteristické pro jednotlivé slovní druhy), tvaroslovného (soubor morfologických kategorií příslušejících lexému) a syntaktického (funkce lexému ve větě) (LVG: 320; MLLVG I: 13); někdy se nevyděluje zvlášť slovotvorné hledisko (VPSV: 437, Kalme–Smiltnece 2001: 10–11), což odpovídá i popisu v českých gramatikách (Štícha et al. 2013: 80–81; Cvrček et al. 2015: 158–160).

Slovní zásoba lotyštiny je rozdělena do deseti základních slovních druhů, shodných s českým systémem,<sup>130</sup> ty pak mohou být dále členěny do dalších specifických subkategorií, z nichž jsou však v LČ zohledněny jen některé, popsané dále. Problematické a hraniční případy, mající často analogie v češtině (Štícha et al. 2013: 83–89), zpracováváme primárně s ohledem na LVG a MLVV jakožto aktuální reprezentativní jazykové zdroje, případně přihlížíme ke korpusovým datům, jsou-li dostatečně průkazná.

Různý slovnědruhový úzus téhož lexému, vzniklý v důsledku konverze a vázaný na určitý kontext (Paegle 2003: 229), je v LČ řešen rozdělením do sa-

---

málo vybraných tvarů, svým způsobem „klíčových“ (Čermák 1995d: 241), přičemž „všechny formy redukce musí však dodržet požadavek jednoznačnosti a také přehlednosti“ (ibid.: 244).

<sup>129</sup> Oboustrannou morfologickou charakteristiku přináší ČLN, ten má však na rozdíl od ČL povahu primárně kódovací (viz též výše 2.2). Oba slovníky, ač počtem hesel (22–23 tis.) obdobné, se co do objemu, respektive formátu podstatně liší.

<sup>130</sup> Jiné je pouze řazení jednotlivých slovních druhů: v lotyšských gramatikách bývají většinou číslovky popisovány před zájmeny, v opačném pořadí občas též spojky s předložkami (LVG: 320; VPSV: 437).

mostatných hesel, odlišených od sebe arabskou číslicí v horním indexu za lemmatem a příslušnou slovnědruhovou značkou, která odpovídá celému heslu, nikoliv jen vybraným významům – tzn. stejně jako jiné druhy homonymie.<sup>131</sup> Lotyšské slovníky, výkladové i překladové, si oddělená lemmata tradičně ponechávají jen pro případy vlastní lexikální homonymie, tj. formální shody bez přímé souvislosti (alespoň v synchronní rovině), dávající přednost výkladu jednotlivých heterofunkčních významů v rámci společného hesla.<sup>132</sup> Nejsou v tom však zcela konsekventní, rozdíly se mohou projevit i v rámci stejného typu lexémů, zvláště je-li slovník plodem práce početnějšího autorského kolektivu.<sup>133</sup>

Okrajovou roli jiného slovnědruhového úzu, naznačenou jeho umístěním až na konec hesla, lze snadno kompenzovat i na úrovni makrostruktury, a sice vhodným uspořádáním homonymních forem v hesláři: ty nejméně obvyklé jsou řazeny úplně nakonec. Frekvenční hledisko může vést k situaci, kdy druhotná, odvozená slovnědruhová charakteristika lexému bude v hesláři předržena původnímu významu (např. některé sekundární předložky vzniklé z příslovcí), tudíž nemusí být jasné, které užití daného lexému je primární; tato informace je však pro uživatele – alespoň tak, jak s ním počítá koncepce LČ – irelevantní.

Polyfunkční tvary lexému jsou tak v LČ pořádány podle pravděpodobnosti, jak se s nimi uživatel setká; obdobně řadíme na základě četnosti i jednotlivé významy v polysémních jednotkách, preferující synchronní zřetel před

---

<sup>131</sup> S tím koresponduje i technická realizace této distinkce v TLexu pomocí atributu Homonymum (2.4.3.3). Názory, zda ve slovnících formálně odlišovat slovnědruhovú (či heterofunkční – Čermák 2010a: 262) homonyma od lexikálních, se různí: Filipec (1995: 23) doporučuje oba typy homonymie lišit v číselném indexu užitím římských, či arabských číslic (případně dělením slovnědruhových homonym v souborné heslové stati speciální značkou); naproti tomu Čermák (1995d: 241), chápající tento případ jako slovnědruhovú polysémii (polyfunkčnost), formální distinkci obou jevů za nezbytnou nepovažuje.

<sup>132</sup> Srov. obdobné stanovisko Vondráčkovu (Štícha et al. 2013: 80). LVG (301) nicméně hovoří o nedostatečné reflexi konverze v lotyšské jazykovědné teorii, na nedůslednosti v chápání homonymie a polysémie upozorňuje též Laua 1969: 64–65.

<sup>133</sup> Což je případ LP, kde jsou např. substantivizované tvary určitých adjektiv zpracovány buď pod samostatným heslem (*kurlais* ‚hluchý [člověk], neslyšící‘), nebo jako jeden z významů pod lemmatem adjektiva v určitém tvaru (*melnais* ‚černoch‘), případně též jako kolokace (*baltais cilvēks* [s markerem *antrop.*] ‚běloch‘) či frazém (*mīkstais* ‚teplouš‘). Výslednou průměrnou různých lexikografických řešení nepřikládáme specifičnosti jednotlivých lexémů jako spíš nejednotnosti popisu a různým přístupům autorů slovníku.

diachronním (srov. příklad *vandālis* v závěru 3.2.4). Samotný způsob zpracování daného významu se nemění, oproti jiným slovníkům se naše řešení liší pouze v jeho vydělení pod samostatné heslo, tedy v upřednostnění vertikálního členění před horizontálním. Důsledná aplikace tohoto pravidla, spolu s řazením podle frekvenčního, nikoliv věcně-logického či historického hlediska, by měla přispět k lepší přehlednosti a většímu uživatelskému komfortu.

Pro účely této práce charakterizujeme po stránce morfologické tři vybrané slovní druhy. Dva z nich jsou plnovýznamové – substantiva a verba, pro lexikografický popis vůbec nejsložitější. Z neplnovýznamových slovních druhů vybíráme konjunkce, specifické mimo jiné tím, že jejich exemplifikace vyžaduje širší, často i nadvětný kontext.

### 3.2.3.1 Gramatická charakteristika substantiv v LČ

Lotyšským substantivům přísluší tři gramatické kategorie, vyjadřované společnou koncovkou: rod (životnost se však na rozdíl od češtiny negramatikalizovala), číslo a pád.

3.2.3.1.1 Dnešní spisovná lotyšтина rozlišuje dva jmenné rody: mužský a ženský,<sup>134</sup> jejich formální distinkce v LČ je nicméně nezbytná pouze u substantiv.<sup>135</sup> Ta označujeme příslušnou zkratkou (m, f, případně m/f pro obourodá), kombinovanou v dolním indexu buď s číslem deklinační třídy (1.–6., viz dále 1.4.1–1.4.6, 0 u nesklonných – 1.4.10), nebo se značkou řadící je k jinému, sekundárnímu deklinačnímu typu (1.4.8–1.4.9).<sup>136</sup> Možné problémy při lexikografickém zpracování představují substantiva obourodá, dále kolísání v rodě a přechýlené lexémy.

---

<sup>134</sup> Někdejší neutrum zaniklo již v předhistorických dobách, substantiva středního rodu přešla k maskulinům a femininům. Dochováno je pouze reliktně v několika málo tvarech zájmen, číslovek a adjektiv (Endzelīns 1951: 395; MLLVG I: 378).

<sup>135</sup> Adjektivní lemmata jsou uváděna výhradně v maskuliní podobě; pronomina většinou také (s výjimkou těch nejfrekventovanějších, u nichž jsou femininí podoby obvykle odkazovým heslem), stejně jako numeralia.

<sup>136</sup> Nakolik je nám známo z jiných slovníků, jde o originální, dosud nikde nepoužité řešení, mající výhodu ve své přímocnosti a logičnosti.

3.2.3.1.1.1 Nevelká skupina lotyšských substantiv je obourodých:<sup>137</sup> táž forma může označovat jak maskulinum, tak femininum – taková označujeme v závislosti na deklinační třídě značkou *m/f<sub>4</sub>* či *m/f<sub>5</sub>*.<sup>138</sup> Obě rodová paradigmata jsou stejná s výjimkou jediného tvaru: dativu singuláru; jako distinktivní jej proto uvádíme v rámci atributu Morfologie, a to v jednotné podobě (maskulinní tvar následovaný tvarem femininním; případně též další formy).

**gnīdļa** 1. *f<sub>4</sub>* (vajīčko vši) *hnida* 2. *m/f<sub>4</sub>* (Dsg m ~am, f ~ai) <hanl> (člověk) *všivák, prašivec, hnida*, veš • g. tu tāds! *ty seš ale všivák! všiváku jeden!*

Zvláštní pozornost je přitom třeba věnovat ženským ekvivalentům, jež mohou inklinovat k nepřirozenosti, umělosti (*ņirga* ‚posměváček × ?‘, *snauga* ‚ospalec, spáč, lenoch × ?‘). V případě možnosti je vhodné od sebe vydělit, byť v rámci téhož významu, český protějšek generického typu od ekvivalentu mužského a ženského:

**aitasgalvja** *m/f<sub>4</sub>* (Dsg m ~am, f ~ai) <hanl> *hlava skopová* || (muž) *hlupák, beran* || (žena) *hlupačka*

3.2.3.1.1.2 Některá lotyšská substantiva, především prefigovaná a kompozita, kolísají v rodě: existují jak v podobě maskulinní, tak femininní. Vzhledem k omezenému rozsahu LČ nepostihujeme varianty nářeční a diachronní, zachycujeme pouze dublety živé v dnešním celouzemním standardním úzu. Významové nuance mezi těmito dubletami bývají často i pro rodilé mluvčí minimální a těžko postižitelné, korpusová data jsou bohužel nedostatečná. Oběma (zřídka třem) tvarům proto odpovídají společné české ekvivalenty, přičemž takovéto případy, následují-li v hesláři bezprostředně po sobě, budou v LČ zpracovány

<sup>137</sup> Jde řádově o desítky lexémů, většinou kolokviální povahy, s expresivním laděním a negativními konotacemi, (LVG: 334), jejichž periferní postavení v lexikonu se odráží i v nízkém zastoupení v LVK. Ačkoliv lotyšský tagger s obourodými substantivy počítá (LVMPK: *kontekstuāla dzimte* ‚kontextový rod‘), je takto v LVK označen jeden jediný lexém. Naopak tvary obourodého lemmatu *paziņa* ‚známý, známá‘ jsou ve všech 213 případech otagovány jako feminina, 7 výskytů Dsg maskulina *paziņam* je chybně lematizováno jako tvar neexistujícího verba *\*pazīnt*.

<sup>138</sup> Kvůli slovům *konferansjē* ‚konferenciér; konferenciérka‘ a *krupjē* ‚krupiér, krupiérka‘ (případně se mohou vyskytnout další) bude třeba zavést též kategorii obourodých nesklonných substantiv. Půjde bezesporu o uzavřenou, málo početnou třídu lexémů, nicméně ze systémového hlediska nelze tuto (potlačit

v rámci jednoho lemmatu:<sup>139</sup> jako první řadíme tvar bezpříznakový či častější (je-li v dvojici takový), případně abecedně předcházející.

**sētmalja** [æ:] f<sub>4</sub>; **sētmajle** [æ:-e] f<sub>5</sub> (Gpl ~ju); **sētmajlis** [æ:] m<sub>2</sub> (~ja) *okolí plotu*

**sprāgo|nis** m<sub>2</sub> (~ŋa) <hov>; **sprāgoŋ|a** m/f<sub>4</sub> (Dsg m ~am, f ~ai) <hov> *mršina, mrcha, zdechlina, chciplotina*

Rodovou dubletu *sāls* ‚sůl‘ zpracováváme na základě jejího popisu v LVG (336–337) a ve výkladových slovnících s přihlédnutím ke korpusovým datům<sup>140</sup> takto:

**sāļs** m<sub>2</sub> (~ls, Dsg ~im, Npl ~ji) / f<sub>6</sub> (~ls, Dsg ~ij, Npl ~lis, Gpl ~ju) **1. sūl** • vārāmā (vārāmais) s.

*kuchyŋská sūl* • ~ls raktuve *solnų dūl* • piebērt ~li [zupai] *přisolit, dosolit [polévku]* [...]

|| **2. <chem> sūl** • skābie / bāziskie ~ji (skābās / bāziskās ~lis) *kyselý / zásadité soli*

• fosforskābes s. *sūl kyseliny fosforečné, fosforečnan* [...] ♦ šīs zemes s. *sūl této země* [...]

*Úzus:* V odborném diskursu (2. význam) převažuje mužská podoba slova, častý je plurál; v neodborném (1.) jsou zase běžné tvary jednotného čísla s ženskými koncovkami.<sup>141</sup>

3.2.3.1.1.3 Přechýlená substantiva jsou v LČ zpracována, na rozdíl od většiny lotyšských překladových slovníků, důsledně jako samostatná lemmata, a to i přes větší nároky na prostor. Proti hnízdování totiž stojí zásadní argument: zdaleka ne každý přechýlený derivát má stejnou sémantickou strukturu jako motivující tvar, lišit se může i kolokabilita obou lexémů. Za základní člen dvojice považujeme výchozí, mužskou formu<sup>142</sup> (fungující nezřídka i jako generické

<sup>139</sup> Formálně v TLexu pomocí elementu Varianta (2.4.4), jenž má stejné atributy jako element Lemma, je však jeho potomkem a sám o sobě existovat nemůže. Srov. ale např. dubletní tvary *apbrīna* (19 lemmat v LVK) × *apbrīns* (15) ‚úžas, obdiv‘, jež musí být zpracovány samostatně (pravděpodobně jedno z nich bude odkazovým heslem), neboť mezi nimi v hesláři figurují přinejmenším další tři lexémy (*apbrīnojami*, *apbrīnojams*, *apbrīnot*).

<sup>140</sup> V LVK je lemmatem *sāls* označeno 568 tokenů, s následnou frekvenční statistikou tvarů: *sāls* 376, *sāli* 166, *sāļu* 11, *sāļus* 6, *sāļiem* 5, *sāļos* 3, *sālim* 1. Srov. kompletní paradigma (bez vokativu, tj. N, G, D, A, I, L): sg. *sāls*, *sāls*, m. *sālim* / f. *sālij*, *sāli*, *sāli*, *sālī*; pl. m. *sāļi* / f. *sālis*, *sāļu*, m. *sāļiem* / f. *sālīm*, m. *sāļus* / f. *sālīs*, m. *sāļiem* / f. *sālīm*, m. *sāļos* / f. *sālīs*. Vytučněny jsou tvary v LVK nedoložené, mimo jiné téměř všechny femininální plurálové formy, což se odráží i v uzuální informaci. (Srov. též LVTK: *sālij* 6, *sālī* 19, *sālīs* 1, *sālīm* 0, *sālīs* 0.)

<sup>141</sup> Srov. řešení v LP i LL, kde jsou tyto dublety považovány za homonyma a zpracovány ve dvou různých heslech, v LL (631–632) jsou obě standardní, v LP (1185) je femininní podoba odkazovým heslem na bezprostředně následující mužskou dubletu, kombinací obou způsobů je pak Andronov 2002.

<sup>142</sup> Až na nečetné výjimky, jako jsou označení profesí tradičně vykonávaných ženami (např. *medmāsa* ‚zdravotní sestra‘ s 45 výskyty v LVK a 562 v LVTK oproti *medbrālis* ‚zdravotní bratr‘ s 0, resp. 3 výskyty; srov. též kolokaci *zdravotní sestra* / *sestřička*, vyskytující se v korpusu SYN2015 854krát, *zdravotní bratr* 12krát) a sociálně podmíněnými reáliemi typu *atraitne* ‚vdova‘.

maskulinum), ta je také zpravidla opatřena bohatší exemplifikací; u ženských protějšků uvádíme jen kolokace frekventované a/nebo příznačné.

**aģen|te** f<sub>5</sub> (Gpl ~šu) **1.** (firmy) *agentka, zástupkyně, jednatelka* || **2.** (tajné služby) *agentka, špionka*

**aģent|s** m<sub>1</sub> **1.** (firmy) *agent, zástupce, jednatel* • kompānijas a. *zástupce (agent) firmy* • tirdzniecības a. *obchodní zástupce (agent)* • apdrošināšanas a. *pojišťovací agent* • nekustamā īpašuma a. *realitní makléř* • biržas a. *burzovní makléř* || **2.** (tajné služby) *agent* • FIB / CIP a. *agent FBI / CIA* • a. *provokators agent provokatér* • savervēt ~u *naverbovat agenta* || *špion, vyzvědač* • atklāt (atmaskot) ~u *odhalit špiona* || **3.** (hybná síla, původce děje) *činitel, agens* || **4.** <chem, med> *čínidlo, agens* • ķīmiskie ~i *chemická čínidla* || **5.** <IT> *agent*

Hnízdování přechýlených lexémů vede kromě jiného k záznamu tvarů marginálních a okazionálních, jež jsou v korpusech doloženy sporadicky či vůbec. Takovéto lexémy nebudou do hesláře LČ zahrnuty. Samotná absence lexému je nositelkou informace o jeho neobvyklé povaze a/nebo řídkém užití.<sup>143</sup> Uživatel by jeho význam měl bez větších potíží odvodit od mužské podoby (např. přechýlené podoby méně frekventovaných obyvatelských názvů typu *urugvajiete*<sup>144</sup> ‚Urugvajka‘, lehce odvoditelného od tvaru *urugvajietis*, jenž v hesláři LČ bude).

3.2.3.1.2 Dnešní lotyština liší dvě čísla: singulár a plurál,<sup>145</sup> stejně jako v češtině tu však existují lexémy s více či méně omezeným číselným paradigmaticem, tedy substantiva mající tvary buď výhradně jednoho čísla (singularia/pluralia tantum), anebo k singuláru, či plurálu tíhnoucí. Od tradičního odlišování těchto dvou skupin lexémů LČ upouští: námi užívaná příslušná značka (*sg*, *pl*<sup>146</sup>) naznačuje, že daný lexém, respektive jeho konkrétní význam se objevuje výlučně či převážně v daném čísle a jeho číselný protějšek je buď úplně nemožný, nebo neobvyklý a málo frekventovaný, případně omezený na spe-

<sup>143</sup> Jakkoliv tu vždy existuje riziko, že si uživatel takovouto absencí mylně vyloží jako opomenutí autorovo.

<sup>144</sup> LVK: lemma *Urugvaja* 6, *urugvajietis* 0, *urugvajiete* 0; LVTK: 130, 9, 2.

<sup>145</sup> Od velmi řídké užívaných reziduí duálu (Endzelīns 1951: 450–460; MLLVG: 382) lze, podobně jako v případě výše zmíněného neutra, odhlédnout. K situaci v litevštině srov. Erhart 1984: 41.

<sup>146</sup> Tedy bez dalšího, limitativního přívlastku: *jen sg.*, *sg tnt.* apod.



cifické oblasti jazyka (básnický jazyk, profesionalismy apod.). V kategorii čísla dochází v lotyštině poslední dobou k výraznému pohybu, zvláště pod vlivem sdělovacích prostředků a běžně mluveného jazyka: roste počet plurálových tvarů substantiv patřících dosud k singulariím tantum a stejně tak se čím dál častěji objevují i singulárové podoby dřívějších pluralií tantum (LVG: 338, 340); obojí lze, byť v omezené míře, najít i v lotyšských korpusech. LČ se pokouší zmíněný vývoj zachytit, a to především cestou méně explicitního značkování lexémů<sup>147</sup> na úkor vhodné exemplifikace; k dispozici je rovněž výslovný popis v rámci gramatické glosy či informace o úzu (viz níže příklad *asins*).

3.2.3.1.2.1 Rozdílné užití čísla bývá mnohdy signálem nového významu (*putekļi* ‚prach‘ × *puteklis* ‚zrnko, částice prachu‘; *jāņogas* ‚rybíz [keř, plody]‘ × *jāņoga* ‚[jednotlivá] bobule rybízu‘), přičemž otázka lexikální příslušnosti daných forem k reprezentativnímu tvaru je problematická, jasná kritéria pro odlišování sporných případů chybějí (LVG: 343). V LČ je tato problematika řešena především v závislosti na míře lexikalizace toho kterého tvaru a také na poměru frekvencí singulárové a plurálové formy, případně se tu projevují i další hlediska jako třeba vzájemná vzdálenost obou lemmat v hesláři. Na základě těchto kritérií vyčleňujeme několik skupin slov, jež zpracováváme obdobně:

3.2.3.1.2.1.1 Tvary vycházející ze stejného základu, avšak plně lexikalizované chápeme jako dvě samostatná lemmata, přičemž plurálové podoby, veskrze pluralia tantum, opatřujeme značkou *pl*.

**rītis** m<sub>1</sub> *rāno*, *jītro* • agrs r. *brzké rāno* • no paša ~a *hned od rāna* • uz ~a pusi; pret ~u *k rānu*, *nad rānem* • no ~a līdz vakaram *od rāna do večera* [...] || ♦ r. esot gudrāks par vakaru *rāno je moudřejší večera* [...]

<sup>147</sup> Neoznačujeme celé skupiny slov pojímané tradičně jako singularia tantum: abstrakta, jména hromadná či látková, ale i vlastní jména aj., zvláště odpovídá-li situace v lotyštině té české; naopak je potřeba upozornit na rozdíly mezi lotyšským a českým lexikonem – a zejména územ.

Srov. též nejednotnost popisu gramatického s lexikografickým v LP (1526): ze 7 lexémů uváděných jako příklad singularií tantum (*piens* ‚mléko‘, *gaļa* ‚maso‘, *nauda* ‚peníze‘, *zelts* ‚zlato‘, *fizika* ‚fyzika‘, *matemātika* ‚matematika‘, *jaunatne* ‚mládež‘) jsou v samotném hesláři značkou *sg. tnt.* označena pouze dvě (*nauda*, *jaunatne*); taktéž ze 7 příkladů pluralií tantum (*durvis* ‚dveře‘, *brilles* ‚brýle‘, *kāpnes* ‚schody‘, *kāzas* ‚svatba‘, *svētki* ‚svátek‘, *milti* ‚mouka‘, *putraimi* ‚kroupy‘) pak nejsou značkou *pl. tnt.* označeno lemma *brilles* a lexém *putraimi* je uveden pod singulárovým lemmatem, s poznámkou *zwl. pl.* [zvláště v plurálu].

**rīt|ji** m<sub>1</sub> pl *východ* • ~u puse *východ*, *východní kraj* • vějš no ~iem *východní vítr*, *vítr od východu*<sup>148</sup>

3.2.3.1.2.1.2 Tvary nelexikalizované, jsoucí navzájem v těsném sémantickém vztahu zpracováváme v rámci jednoho lemmatu, s jasným vydělením jednotlivých významů a uvedením čísla pro ně typického.

**zir|ŋi** m<sub>2</sub> 1. pl *hrách* • ~ŋu zupa / biezenis *hrachová polévka* / *kaše* • pelēkie z. ar speķi *šedý hrách se špekem* (tradiční lot. pokrm) • ~ŋu pāksts *hrachový lusk* • sēt / lobīt ~ŋus *sít (zasévat) / loupat hrách* [...] || sg (~nis, Gsg ~ŋa) <bot> *hrách* (Pisum) • sējas ~nis *hrách setý* (Pisum sativum) || 2. sg (~nis, Gsg ~ŋa) *hrášek*, *zrníčko hrachu* • ~ŋa lielumā *velikosti hrachu* || ♦ pliks *kā ~nis nahý jako prst (Adam)* ♦ bērt [vārdus] *kā ~ŋus jako když bičem mrská*

**zi|vs** f<sub>6</sub> (~vs, Gpl ~vju) 1. *ryba* • salsūdens / jūras ~vis *sladkovodní / mořské ryby* • ~vju zupa / eļļa / konservi *rybí polévka / olej / konzerva* • ceptas ~vis *pečené ryby* • ~vju pirkstiņi *rybí prsty* • dziļūdens ~vis <zool> *hlubinné ryby* [...] || 2. pl Zi|vis (Gpl ~vju) <astr> *Ryby* • viņš ir dzimis ~vju *zímě narodil se ve znamení Ryb* || ♦ *kā z. ūdenī jako ryba ve vodě* ♦ liela z. *velká ryba, velké zvíře* [...] <sup>149</sup>

Výjimku tvoří případy, kdy jsou oba tvary od sebe v hesláři vzdálené a identifikace lemmatu by mohla uživateli činit potíže: Tak jsou např. hesla *rīss* a *rīsi*, oddělená od sebe v hesláři cca desítkou lemmat, zpracována samostatně, třebaže sémanticky jsou oba tvary ve stejném poměru (rostlina – její plod) jako jiné dvojice, jež by v hesláři po sobě bezprostředně následovaly.

3.2.3.1.2.1.3 Číselné dublety substantiv *asins* × *asinis* ‚krev‘, *debess* × *debesis* ‚nebe, obloha‘, *smilts* × *smiltis* ‚písek‘ zpracováváme pod singulárním lemmatem,<sup>150</sup> plurálovou formu uvádíme jako odkazové heslo pouze v případě jejich přílišné vzájemné vzdálenosti v hesláři.

**asi|ns** f<sub>6</sub> (~ns, Gpl ~ŋu) *krev* • izliet ~nis (*par ko*) *prolít krev* (*za koho/co*) • nodot ~nis *darovat krev* • ~ns lāse *kapka krve* • ~nīm piepludušas *acis oči podlité krví* • Kristus *miesa un ~nis <náb> tělo a krev Páně* [...] || <anat, med> (v odborném diskurzu převažuje singulár) • ~ns analīze *rozběr krve* • ~ns saindēšanās *otrava krve* • ~ns pārliešana *transfúze krve* • ~ns noņemšana *odběr*

<sup>148</sup> Homonymie *rīti*<sup>1</sup> (plurale tantum) ‚východ‘ × *rīti*<sup>2</sup> (plurálový tvar lexému *rīts*) ‚rána, jitra‘ už však kvůli omezenému prostoru zachycena není, druhý tvar je implicitně zahrnut v příslušném lemmatu; v zásadě by tomu však nic nebránilo, zvláště u frekventovaných homonym. Tradiční lotyšská lexikografie běžně postihuje homonymii pouze na rovině lemmatické; k jiným rovinám viz Grīsele 2008.

<sup>149</sup> Srov. též heslo *ryba* v ČLN (253): **rybja** f, (-y) 1 *zi|vs f*, (Gpl ~vju); chov ryb *zivkopība f*, lov ryb *zvejja f*, (na udici) *makšķerēšana f*, chytat ~y *zvejlot (II.)*, (na udici) *makšķerēt (II.)*; chodit na ~y *iet makšķerēt* ♦ *zdravý jako r. vesels kā rutks*; němý jako r. *mēms kā zivs*; být jako r. *ve vodē justies kā zivij ūdenī*; ani r., ani rak *ne cepts, ne vārīts* 2: **R~y** <astr> *Zi|vis f*, pl, (G ~vju)

<sup>150</sup> Srov. nejasnosti v lemmatizaci těchto lexémů v LVK (*asins* 510, *asinis* 250, *debess* 736, *debesis* 15, *smilts* 550, *smiltis* 0), ale rovněž různá pojetí v lotyšských slovnících.

*krve* • ~s grupa krevní skupina • ~ns aina krevní obraz • arteriālās / venozās ~nis arteriālnī (tepennā) / venóznī (žilnī) krev • ~ns sastāvs skladba krve • ~ns vēzis rakovina krve, leukémie [...] || ♦ sarkans kā ~nis červēnī (rudī) jako krev ♦ ~nis sakāpj (saskrien) galvā krev stoupá do hlavy ♦ ~nis sastingst dzīslās krev tuhne v žilách ♦ viņai dzīslās rit zilās ~nis koluje jí v žilách modrá krev ♦ gēnos viņam dažādu tautu ~nis v genech má krev různých národů [...]

3.2.3.1.2.1.4 V lotyšských zejména výkladových slovnících bývají za reprezentativní tvary považovány plurálové formy lexémů označujících různé skupiny osob, kromě obyvatelských jmen též např. názvy příslušníků mnišských řádů (*cistercieši* ‚cisterciáci‘) či náboženských vyznání (*krišnaīti* ‚křišnovci‘), filosofických škol (*eleāti* / *elejieši* ‚eleaté‘), politických či kulturních uskupení (*veclatvieši* ‚starolotyši‘) aj. Oproti této praxi preferuje LČ singulárové podoby lemmat před plurálovými, přičemž přechýlené tvary nehnízduje, ale v případě dostatečné frekvence je uvádí jako samostatná hesla (1.1.3). Plurálovou formu, zahrnovanou některými překladovými slovníky nepřesně pod kolokace, pro její snadnou odvoditelnost (*kriev/s* ‚Rus‘ – *~i* ‚Rusové‘) neuvádíme.

**kriev|s** *m*<sub>1</sub> *Rus* • ~u valoda *ruština* • ~u alfabēts *azbuka* • ~u filoloģija *ruská filologie*, *rusistika* • ~u laiki <hov> *sovětské časy*, *sovětská doba* (období Lotyšské SSR: 1940–91) [...] || ♦ ~u rulete *ruská ruleta* ♦ iet / iesaukt ~os <hov, sov> *narukovat* / *povolat do armády*

**krievie|te** *f*<sub>5</sub> (Gpl ~šu) *Ruska*

**amī|tis** *m*<sub>2</sub> (~ša) <slang> *Amík*, *Amerikán*

**krišnaī|s** *m*<sub>1</sub> *krišnovec*, *krišňák*

**rozenkreicer|s** [o] *m*<sub>1</sub> **rozenkreicie|tis** [o] *m*<sub>2</sub> (~ša) *rosekrucián*

**veclatvie|tis** *m*<sub>2</sub> (~ša) **1.** <hist> *starolotyš* (kulturní pracovník v Lotyšsku v 50.–60. l. 19. stol.) [...]

3.2.3.1.2.1.5 Nekonzistence panuje i v lexikografickém popisu některých botanických a zoologických termínů.<sup>151</sup> U nich kromě slovníků přihlížíme rovněž k encyklopedickým příručkám: je-li daný taxon doložen v singulárové podobě, přebíráme ji jakožto lemma. Plurálové formy omezujeme pouze na případy typu *segsēkļi* ‚krytosemenné‘, *bezmugurkaulnieki* ‚bezobratlí‘ apod.

<sup>151</sup> Viz např. v MLVV heslo *ašķi* ‚přeslička‘ (singulárovou podobu *ašķis* systém nevyhledá, lze ji však najít v LP, LL i LR – a upřednostní ji i LČ), odkazující k synonymnímu výrazu (rovněž plurálovému) *kosas*, přičemž heslář obsahuje pouze singulárový tvar *kosa*.

3.2.3.1.2.1.6 U dalších problematických lexémů je třeba přihlížet k jejich frekvenci a přirozenosti jak ve výchozím, tak cílovém jazyce: singulárová podoba *dvīnis* ‚dvojče‘ je v lotyšských korpusech dostatečně doložena, byť dominuje plurálový tvar,<sup>152</sup> ne tak už tvar *trīnis* ‚trojče‘, *četrītis* ‚čtyřče‘, *piecītis* ‚paterče‘<sup>153</sup> atd. Singulárovou formu lze uvést u plurálových lemmat v rámci atributu Morfologie, eventuálně to může být i alternativní řešení v některých výše uvedených případech, kde výrazně převažuje plurálový úzus.

**četrī|ši** m<sub>2</sub> pl (sg. ~tis, Gsg ~ša) *čtyřčata*

**bezmugurkaulnieki** m<sub>1</sub> pl (sg. ~s) <zo<sub>ol</sub>> *bezobratlí* (Invertebrata)

**veclatvie|ši** m<sub>2</sub> pl (sg. ~tis, Gsg ~ša) 1. <hist> *starolotyši* (kulturní pracovníci v Lotyšsku v 50.–60. l. 19. stol.) [...]

**cistercie|ši** m<sub>2</sub> pl (sg. ~tis, Gsg ~ša) <cír< cisterciáci • ~šu ordenis *cisterciácký řád*

**eleāt|i** m<sub>1</sub> pl (sg. ~s) <filoz> *eleaté*<sup>154</sup>

3.2.3.1.3 LČ se drží tradičního pojetí morfologicko-syntaktické kategorie pádu, jak je prezentován ve většině lotyšských gramatik i gramatických přehledů v překladových slovnících: pádový systém má stejně jako v češtině sedm pádů.<sup>155</sup> Liší se pouze jejich označení a řazení; lotyšské pády jsou nazývány latin-

<sup>152</sup> V LVK je lexém *dvīnis* (též jeho přechýlená podoba *dvīne* a oba deminutivní deriváty *dvīnītis/dvīnīte*) lemmatizován jako singulár; singulárových tvarů je nicméně 6 (*dvīne* 2) oproti 65 (7) plurálovým (*dvīnītis* 1 oproti 4, *dvīnīte* 0 oproti 1), v LVTK min. 129 výskytů sg. podoby *dvīnis*, min. 7 *dvīnīte*; sg. tvary deminutivních derivátů v řádech jednotek.

<sup>153</sup> Nutno však brát v potaz též polysémii: lexém *piecītis* v kolokviálním významu ‚pětieurová bankovka, (dříve pětilatová, pětirublová – viz též LVSV: 357), bůr‘ má v LVK 7 výskytů oproti 6 *piecīši* ‚paterčata‘ (z toho 5 v kolokaci *Čikāgas piecīši* [název hudební skupiny]), v LVTK se objevuje min. v 70 výskytech z 207. Nadto LVSV (ibid.) uvádí ještě jeden význam (‚známka ve škole, odpovídající v české klasifikaci jedničce‘) a doklady v LVTK svědčí o dalším významu, synonymním s výrazem *piecnieks* ‚pětice [zvláště lidí]‘.

<sup>154</sup> Srov. NASCS (136, 201), kde jsou oba lexémy uvedeny pod singulárovým lemmatem: *cisterciāk*, *eleat*. Korpusevá data hovoří jednoznačně ve prospěch plurálové formy: lemma *cisterciāk* v SYN2015 47× v pl., 4× v sg.; *cisterciačka* 9× v pl., 0× v sg., 2× tvar *cisterciāni* a 1× *cisterciāti*; lemma *eleat* v SYN2015 nedostatečná data, v SYN (pod nesprávným lemmatem *eleata/eleatū*) výhradně v sg., a to 17×. V LVK se lexém *cistercietis* vyskytuje jen jednou – v plurálu, v LVTK 129× v pl., 6× v sg., *cisterciete* 5× v pl., 1× v sg.; lexém *eleāts* ani *elejietis* se v lotyšských korpusech neobjevuje ani v jednom z čísel.

<sup>155</sup> Nejčastěji bývá zpochybňován samostatný status instrumentálu, jenž v důsledku jazykového vývoje formálně splýnul v singuláru s akuzativem a v plurálu s dativem; od těchto pádů se liší pouze spojením s předložkou *ar* ‚s‘ – ta se však často vynechává, zvláště v běžně mluveném jazyce a zejména v plurálu. Čistě formálně pohlíží na instrumentál např. lotyšský tagger SemTi-Kamols, jenž všechny jeho výskyty přiřazuje buď k akuzativu, nebo dativu, což mimo jiné znemožňuje zjistit frekvenční distribuci pádových forem v lotyštině na základě korpusevých dat. LČ se kloní k tradičnímu pojetí instrumentálu jako samostatného gramému sice bez vlastní pádové formy, avšak se specifickou sémantikou odlišnou od ostatních pádů: vyjadřuje především adverbiální významy (prostředek, nástroj, sdílení děje, čas, způsob či míru). Jeho systémová pádová homonymie je vyjádřena pomocí lomítka: A/Isg, D/Ipl (viz též LVG: 353).

skými názvy, nikoliv čísla a uvádějí se konvenčně v tomto pořadí: *nominatīvs* (nominativ, N), *genitīvs* (genitiv, G), *datīvs* (dativ, D), *akuzatīvs* (akuzativ, A), *instrumentālis* (instrumentál, I), *lokatīvs* (lokativ, L), *vokatīvs* (vokativ, V). V LČ jsou zastoupeny zkratkou, respektive kombinací zkratek pádu a čísla (např. Dsg, Vsg, Gpl, A/Isg, D/Ipl). Z deklinačního paradigmatu vybíráme jen některé pádové formy, zvláště ty podléhající hláskové změně,<sup>156</sup> nepravidelné či dubletní, ostatní tvary jsou od nich odvoditelné (a verifikovatelné v příslušném vzorovém paradigmatu v přehledu gramatiky na konci slovníku). Klíčovým je pro nás genitiv singuláru;<sup>157</sup> je-li uváděn, pak vždy na prvním místě (logicky následuje po nominativu singuláru, tj. lemmatu), a tudíž bez příslušné zkratky.<sup>158</sup> Naopak pádové tvary neuvádíme, je-li paradigma pravidelné a nevykazuje-li žádné hláskové alternace (s výjimkami popsány níže); k jeho rekonstrukci plně postačí patřičná rodově-deklinační značka.

**ozol|s** m<sub>1</sub> <bot> *dub* (Quercus) [...]

**vasar|a** f<sub>4</sub> *léto* [...]

**brā|lis** m<sub>2</sub> (~|a, Vsg ~l!/~lī!) **1.** (sourozenec) *bratr* [...]

**al|us** m<sub>3</sub> (~us) *pivo* [...]

**mā|te** f<sub>5</sub> (Vsg ~t!/~te!, Gpl ~šu) **1.** (rodička) *matka* [...]

**krā|sns** f<sub>6</sub> (~sns, Gpl ~šņu) *pec, kamna* [...]

**miegami|ce** m/f<sub>5</sub> (Dsg m ~cem, f ~cei; Gpl ~ču) (muž i žena) *ospalec, spáč* [...]

Podobně vyčleňují některé přístupy z deklinačního paradigmatu také vokativ, ani on nemá vlastní pádovou koncovku, od ostatních pádů se odlišuje i z hlediska syntaktického. I zde se přikláníme k tradičnímu pojetí, jež přiznává vokativu samostatnou povahu. Oproti jiným překladovým slovníkům, jež tento pád přehlíží (zahrnují jej pouze do paradigmat vzorových substantiv v přehledu gramatiky), bude tvar vokativu soustavně uváděn u vybraných skupin lexémů, kde je jeho užití logické a očekávatelné, mj. u všech křestních jmen, vybraných označení osob, ale např. i nadávek apod.

<sup>156</sup> Proti uvádění pravidelných hláskových alternací je Andronov (1998: 142; 2002: 9). Viz též následující poznámku.

<sup>157</sup> Naše rozhodnutí je ovlivněno lotyšskou lexikografickou tradicí, jakkoliv má větší distinkční platnost Dsg, jenž je pro každou jednotlivou deklinaci svěbytný a jako jediný netvoří homonymní formy (LVG: 357; Andronov 1998, 2002: 9, 332). Gsg u některých maskulin 2. deklinace však vykazuje hláskovou změnu, na niž je vhodné uživatele upozornit. (Srov. **níže** uvedený příklad *brālis* s řešením Andronovovým **brālis** *ōpam*, kde je hlásková alternace naznačena jedinou kolokací *brāļu kapi*; avšak např. minimalisticky pojaté heslo *cālis* *цятлэнок*, spolu s mnoha dalšími, postrádá jakoukoliv gramatickou charakteristiku, jež je podle autora [Andronov 1998: 145] nadbytečná.)

<sup>158</sup> „[...] (právě absence označení jednoho typu, obv. nejčastějšího, mezi ostatními typy označenými může být sama signálem).“ (Čermák 1995d: 243).

Sémantika jednotlivých pádů je slovníkem na rozdíl od gramatického popisu zachytitelná pouze v omezené míře, proto se jí tu na tomto místě nevěnujeme. Neznamená to však rezignaci na pokus naznačit ji v LČ vhodně volenými příklady (případně glosou), vedle toho by této problematice měla být věnována pozornost v gramatickém přehledu na konci slovníku – jednotlivé pády tu budou pojaty kontrastivně, s důrazem na jejich odlišnou sémantiku v lotyštině a češtině.

3.2.3.1.4 Pro klasifikaci skloňování lotyšských substantiv existuje vícero kritérií, jejich přehled uvádí LVG (357). Vzhledem k synchronní povaze LČ upouštíme od kritérií diachronních a v souladu s lingvistickou tradicí zavedenou začátkem druhé poloviny 20. století (ibid.) rozlišujeme deklinační typy na základě rodu a koncovky Nsg daného substantiva; takto bývá rozlišováno šest deklinačních tříd, označovaných číslem, případně i historickým kmenem. V LČ je číslo deklinační třídy kombinováno s rodovou zkratkou. Substantiva nepatřící k žádné z deklinačních tříd jsou pojednána dále.

3.2.3.1.4.1 K první, početně nejobsáhlejší<sup>159</sup> deklinaci patří maskulina končící v Nsg na –s nebo –š, v Dsg na –am. V paradigmatu nedochází k žádným hláskovým změnám, není proto zapotřebí uvádět žádný tvar s výjimkou zvláštních forem vokativu (s nulovou koncovkou, případně forem dubletních):

**skolotāj|s** m<sub>1</sub> (Vsg ~!) *učitel* [...]

**Aivar|s** m<sub>1</sub> (Vsg ~!) lot. mužské křestní jméno

**tēv|s** [æ:] m<sub>1</sub> (Vsg ~s!/~!) **1.** (rodič) *otec* [...]

3.2.3.1.4.2 K druhé deklinaci náleží maskulina končících v Nsg na –is a také několik maskulin končících na –s; společná je jim koncovka Dsg –im. K hláskové změně dochází u některých z nich v Gsg a v plurálových formách, na tu explicitně upozorňujeme:

---

<sup>159</sup> Frekvenční statistika substantiv v LVK podle příslušnosti k deklinační třídě: 1. 688 086; 2. 131 821; 3. 7640; 4. 586 290; 5. 224 210; 6. 35 233; 0. [nesklonná substantiva] 3470; r [reflexivní substantiva] 9510.

**divda|bis**<sup>1</sup> m<sub>2</sub> (~bja) <lingv> *příčestí, participium* [...]

**cā|lis** m<sub>2</sub> (~|a) *kuře* [...]

**sliek|snis** m<sub>2</sub> (~šŋa) **1. práh** [...]

Ve snaze předejít uživatelově nejistotě uvádíme tvar Gsg rovněž u lexémů, kde tato hlásková alternace neproběhla, ač jí ze systémového hlediska nic nebrání,<sup>160</sup> případně pokud existují konkurenční dublety:

**vie|sis** m<sub>2</sub> (~sa, Vsg ~si!) *host, návštěvník* [...]

**gaišma|tis** m<sub>2</sub> (~ta) *blondýn, blond'ák* [...]

**Ga|tis** m<sub>1</sub> (~ta, Vsg ~ti!) lot. mužské křestní jméno

**Visval|dis** m<sub>1</sub> (~ža/~da, Vsg ~di!) lot. mužské křestní jméno

Gsg neuvádíme pouze u lexémů, jejichž kmen končí na měkký konsonant, a další hlásková změna tak není možná. Případná kolmička se oproti výše uvedeným případům posouvá o grafém doprava, oddělujíc kmen od koncovky:

**ka|jis** m<sub>2</sub> *kočka* [...] || <zool> (Felis) [...]

**ķīmiķ|jis** m<sub>2</sub> *chemik* [...]

**Uģ|jis** m<sub>2</sub> (Vsg ~i!) lot. mužské křestní jméno (≡ č. Hugo)

U uzavřené množiny maskulin *akmens* ‚kámen‘, *asmens* ‚ostří, břit‘, *mēness* ‚měsíc [družice Země]‘, *rudens* ‚podzim‘, *ūdens* ‚voda‘, *zibens* ‚meč‘, řazené tradičně k 2. deklinaci, dochází k hláskové alternaci pouze v plurálových tvarech. Na tuto výjimku upozorňujeme uvedením tvarů Gsg (shodný s Nsg, v běžně mluveném jazyce se však užívá rovněž analogické palatalizované formy) a Npl:

**akme|ns** m<sub>2</sub> (~ns/<hov> ~ŋa, Npl ~ŋi) *kámen* [...]

**rude|ns** m<sub>2</sub> (~ns / <hov> ~ŋa, Npl ~ŋi) *podzim* [...]

Substantivum *suns* ‚pes‘ (původně \**sunis*) má tvar Gsg *suŋa*, v běžně mluveném jazyce pak spisovnému tvaru Dsg *sunim* konkuruje podoba *suŋam*:

<sup>160</sup> LVG (360) tuto tendenci vysvětluje „snahou vyhnout se nelibozvučnosti tvaru“ a, zvláště u dvouslabičných osobních jmen, přílišným odklonem „od fonetického složení základního tvaru“ (*Atis* – *Aša*; *Rūdis* – *Rūža*), což může být mluvčími nelibě vnímáno. (Srov. též ibid.: 137.)

**su|ns** m<sub>2</sub> (~ŋa, Dsg ~nim / <hov> ~ŋam, Npl ~ŋi) *pes* [...] || <Zool> (Canis) [...]

Tyto hovorové, nekodifikované dublety, pronikající i do psaných textů a neřídce zastoupené v lotyšských korpusech,<sup>161</sup> uvádíme v LČ jako příznak možných jazykových změn; LVG (315) hovoří o znacích smíšené deklinace, tu však lotyšská lexikografická díla dosud nereflektovala.

3.2.3.1.4.3 Ke třetí deklinaci patří maskulina zakončená v Nsg na *–us*, v Dsg na *–um*. Z pádových tvarů uvádíme Gsg, totožný s Nsg, případně též Vsg:

**tirg|us** m<sub>3</sub> (~us) 1. *trh, tržiště* [...]

**Sald|us** m<sub>3</sub> (~us) město v záp. Lotyšsku; správní středisko stejnojmenného kraje

**Mik|us** m<sub>3</sub> (~us, Vsg ~u!) lot. mužské křestní jméno (≡ č. Michal)

Počet substantiv spadajících do této deklinační třídy je omezený na pár desítek<sup>162</sup> a dále neroste, naopak některá maskulina odtud v minulosti přešla do první deklinace. V LVG (362) uváděnou trojici příkladů *cirkus* ‚cirkus‘, *krogus* ‚hospoda, hostinec‘, *strīdus* ‚spor, pře‘ zahrnujeme rovněž do hesláře LČ jako odkazová hesla. Obdobně je zpracována i trojice feminin *dzirnis* ‚mlýn‘, *pelus* ‚plevy‘, *ragus* ‚saně‘ patřících historicky k u-kmenům, jež se nově skloňují podle 4. deklinace. Jejich užití je dnes omezeno dialektálně, ale lze se s nimi setkat též ve folklorních a uměleckých textech (LVG: 361). Odkazový charakter hesla nevylučuje uvést za odkazem specifickou kolokabilitu a zejména frekventované frazémy.

**krog|us** m<sub>3</sub> (~us) ↑ krogš ♦ ~us brāļi *bratři z mokré čtvrti*

**dzirn|us** f<sub>3</sub> pl (Gpl ~u) <zast, dial> ↑ dzirnavas

<sup>161</sup> Tvar *akmeņa* v LVK 4 výskytů, v LVTK 217 výskytů (+ 19 bez diakritiky); *asmeņa* 2/71; *rudeņa* 3/12+2, *ūdeņa* 2/43+7, *zibeņa* 3/15; tvar *mēneša* (382/8794+253) je problematický kvůli homonymii s Gsg *mēnesis* ‚měsíc [časová jednotka]‘ a zasloužil by si samostatnou studii.

<sup>162</sup> Dotaz [tag="N...3\*"] najde v LVK celkem 7640 tokenů v 331 typech, z nichž však většina s frekvencí ≤ 10 (306 lemmat) je chybně lemmatizovaných, nejčastěji jde o lexémy cizího původu (*Endr[j]ū – Endr[j]us*, *tīū* [interjekce] – *tīus*, *Haus*, *Oknarium – oknarius* aj.) Frekvenční špička u-kmenů podle LVK (tvořící téměř 90 % všech výskytů): *tirgus* 2737, *vidus* 1523, *lietus* 1234, *alus* 456, *medus* 148, *klepus* 72, *apvidus* 70, *viltus* 62, *premius* 58, *gadatirgus* 58, z nichž pouze předposlední je špatně lemmatizováno (*Premium*).



3.2.3.1.4.4 Ke čtvrté deklinační třídě patří feminina zakončená v Nsg na *–a*, v Dsg na *–ai*, dále maskulina končící v Nsg na *–a*, v Dsg na *–am* a rovněž nepočetná skupina obourodých substantiv lišících se v paradigmatu právě ve tvaru Dsg (o nich viz výše 1.1.1 i s ukázkou hesel). Feminina nevykazují žádnou hláskovou změnu, obdobně jako u lexémů z první deklinace uvádíme proto jen případný zvláštní tvar Vsg:

**dien|a** f<sub>4</sub> *den* [...]

**skolotāj|a** f<sub>4</sub> (Vsg ~!) *učitelka* [...]

**mās|a** f<sub>4</sub> (Vsg ~a! / ~!) **1.** (sourozenec) *sestra* [...]

U maskulin upozorňujeme na specifický tvar Dsg, případně i na Vsg:

**puik|a** m<sub>4</sub> (Dsg ~am, Vsg ~a!) *kluk, chlapec, hoch* [...]

**dalailam|a** m<sub>4</sub> (Dsg ~am) *dalajláma* [...]

3.2.3.1.4.5 Do páté deklinační třídy jsou řazena feminina zakončená v Nsg na *–e*, v Dsg na *–ei* spolu s nečetnou skupinou obourodých substantiv,<sup>163</sup> lišících se pouze v Dsg: maskulina mají koncovku *–em*, feminina *–ei*. Jedinou pádovou formou, kde může dojít k hláskové alternaci, je Gpl; z důvodu velkého množství slov (většinou cizího původu), jejichž kmen se nemění, uvádíme však tento tvar soustavně u všech lexémů patřících k této deklinaci.

**apro|ce** f<sub>5</sub> (Gpl ~ču) **1.** *náramek* [...]

**miegami|ce** m/f<sub>5</sub> (Dsg m ~cem, f ~cei; Gpl ~ču) (muž i žena) *ospalec, spáček* [...]

**dzī|ve** f<sub>5</sub> (Gpl ~vju) **1.** *život* [...]

**zvaig|zne** f<sub>5</sub> (Gpl ~žņu) **1.** *hvězda* [...]

**sacīk|stes** f<sub>5</sub> pl (Gpl ~šu) **1.** *soutěž, závod* [...]

**ķē|de** f<sub>5</sub> (Gpl ~žu) **1.** *řetěz* [...]

**brīnumsvecī|te** f<sub>5</sub> (Gpl ~šu) *prskavka*

**dentis|te** f<sub>5</sub> (Gpl ~tu) *zubařka, dentistka, stomatoložka* [...]

<sup>163</sup> Protože obourodá substantiva nejsou otagována, je třeba hledat maskulina zakončená na *–e*. Dotaz [tag="N.m.\*"&lemma="\*.e"] najde 201 výskytů, jde převážně o osobní jména, zbytek jsou chybně označované lexémy.

**logopē|de** f<sub>5</sub> (Gpl ~du) *logopedka* [...]

**gaišma|te** f<sub>5</sub> (Gpl ~tu) *blondýna, blondýnka* [...]

Má-li lexém dubletní tvary Gpl, uvádíme obě podoby:

**cunf|te** f<sub>5</sub> (Gpl ~šu / ~tu) <hist> *cech* [...]

**zebiek|ste** f<sub>5</sub> (Gpl ~šu / ~stu) <zool> *lasice* (Mustela) [...]

Mimoto zaznamenáváme i případnou dubletu Vsg:

**mā|te** f<sub>5</sub> (Vsg ~t!/~te!, Gpl ~šu) **1.** (rodička) *matka* [...]

**Kristīn|e** f<sub>5</sub> (Vsg ~e!/~!) lot. ženské křestní jméno (≡ č. Kristýna)

3.2.3.1.4.6 K poslední, šesté deklinační třídě patří feminina končící v Nsg na *–s*, v Dsg na *–ij*. Z pádových forem jsou v LČ uvedeny Gsg (shodný s Nsg) a Gpl, kde ve většině případů dochází k hláskové alternaci, jediné v celém paradigmatu:

**ugu|ns** f<sub>6</sub> (~ns, Gpl ~ŋu) **1.** *oheň* [...]

**sir|ds** f<sub>6</sub> (~ds, Gpl ~žu) **1.** *srdce* [...]

Aby uživatel nezůstal na pochybách, uvádíme i zde – obdobně jako u předešlé deklinace – tvar Gpl systematicky, tzn. i v případech, kdy k palatalizaci nedošlo:

**dak|ts** f<sub>6</sub> (~ts, Gpl ~tu) *knot* [...]

**acs** f<sub>6</sub> (acs, Gpl acu) **1.** (orgán zraku) *oko* [...]

**broka|stis** f<sub>6</sub> pl (Gpl ~stu) *snídaně* [...]

3.2.3.1.4.7 Podle výše popsaných šesti deklinačních typů se skloňuje většina lotyšských substantiv, vedle toho existují i sekundární způsoby deklinace (LVG 2015, 366nn.). Zvláštní paradigma mají reflexivní substantiva, tvořená od infinitivního kmene reflexivního slovesa pomocí formantu *–šanās*. Paradigma je neúplné, zahrnujíc pouze dva tvary: jeden (*–šanās*) se používá v Nsg, Gsg, Npl a Apl, druhý (*–šanos*) v A/Isg a Gpl, přičemž tvary Dsg, D/Ipl a obě lokativní

formy zůstávají neobsazené. Užití těchto lexémů je proto syntakticky značně omezeno a je typické zvláště pro singulár (LVG 2015: 366); v hesláři LČ je obsažena pouze frekvenční špička těchto lexémů<sup>164</sup> a k jejich vyčlenění ze skupiny ostatních feminin se používá značka  $f_{refl}$ .

**vēlēšan|ās**  $f_{refl}$  (~ās, A/I sg, Gpl ~os) **1. přání, touha** [...]

**precēšan|ās**  $f_{refl}$  (~ās, A/I sg, Gpl ~os) (muž) **ženitba**; (žena) **vdavky** [...]

3.2.3.1.4.8 Podle složené adjektivní deklinace se skloňují substantiva vzniklá slovnědruhovou konverzí z adjektiv a ohebných participií. V závislosti na míře jejich lexikalizace a také na slovnědruhový úzus jejich českého ekvivalentu jsou zařazena do LČ jako samostatná hesla; opatřena jsou speciální značkou  $m_{adj}$ , respektive  $m_{partic}$  bez další morfologické charakteristiky:

**balt|ais**  $m_{adj}$  *bēloch* [...]

**mīkst|ais**  $m_{adj}$  <hanl> *teplouš, buzerant, buzna*

**ēdam|ais**<sup>165</sup> [æ:]  $m_{partic}$  <hov> *jídlo* • ē. un dzeramais *jídlo a pití* • garšīgs ē. *dobré jídlo* • sagādāt [kaut ko] ~o *obstarat něco k snědku* • pilns galds ar ~ajiem *plný stůl jídla*

3.2.3.1.4.9 Substantiva cizího původu nespádající koncovkou Nsg do žádné z deklinačních tříd zůstávají nesklonná,<sup>166</sup> vyskytují se tak v jednom jediném tvaru. Takovéto lexémy označujeme příslušným rodem v kombinaci s číslicí 0 v dolním indexu, tedy analogicky k číslům deklinačních tříd.

**eiro** [o]  $m_0$  *euro* [...]

**kanoe** [o·e]  $f_0$  *kánoe* [...]

**Dili**  $f_0$  *Dillí*

<sup>164</sup> Dotaz [tag="N...r"] nalezne v LVK 888 různých lemmat (celkem 9510 výskytů).

<sup>165</sup> Jiné je **ēdam|s** [æ:]  $m_{partic}$  *jedlý, požitelný* [...], tedy participiální adjektivní forma, jež může mít podobu jak neurčitou (*ēdamas ir tikai sēnes cepurītes* 'jedlé jsou pouze kloboučky hub'), tak určitou (*ēdamās sēnes* 'jedlé houby'), reprezentativní tvar, pod nímž bývá zařazen do hesláře, je však až na nemnoho výjimek neurčitý a vždy maskulinní, stejně jako v případě substantivizovaného participia.

Obdobně např. token *labais* může mít trojí platnost, ačkoliv formálně jde vždy o určitý tvar adjektiva *labs*: buď je to tvar nelexikalizovaný ('dobrý'), nebo lexikalizovaný ('pravý [opak levý]'), anebo konverzí nabyvší úzus substantiva (<filoz> 'dobro'), přičemž druhý a třetí tvar mají vlastní lemma (tj. homonymní *labais*<sup>1</sup> a *labais*<sup>2</sup>), kdežto první tvar je začleněn pod lemma *labs*.

<sup>166</sup> Dotaz [tag="N...0"] najde v LVK 3470 výskytů, lemmatizovaných do 109 různých typů, nejfrekventovanějšími jsou: *auto* 544, *eiro* 522, *radio* 474, *kino* 355, *Oto* 205, *foto* 151, *video* 135, *kazino* 118 a *dinamo* 85.

Případné nesklonné dublety uvádíme jako variantu lemmatu, tou může eventuálně být i naopak sklonná forma.

**pianīns** <sub>m<sub>1</sub></sub>; **pianīno** [o] <sub>m<sub>0</sub></sub> *pianino*  
**pueblo** [-'e-o] <sub>f<sub>0</sub></sub>; **puebla** [-'e-o] <sub>f<sub>4</sub></sub> *pueblo*

3.2.3.1.4.10 Specifickou skupinu nesklonných lexémů, vydělovanou v lotyšské lingvistice od 70. let minulého století, představují tzv. *ģenitīvēņi* (LVG: 368–369; Nītiņa 1994).<sup>167</sup> Jde o ustrnulé tvary Gsg či Gpl vystupující ve funkci přívlastku:<sup>168</sup> označují vlastnost či příznak rozvíjeného substantiva. Ze slovotvorného hlediska jsou to nejčastěji složeniny nebo prefigované deriváty. Jejich těsné spojení s rozvíjeným členem naznačují mnohé slovníky pomocí tiretu či pomlky za daným lexémem (např. *bezmaksas-*), v LČ tento způsob zápisu vyhraujeme pro prefixy a první části složenin (viz 3.1.2.4.2) a ke klasifikaci lexikalizovaných genitivů užíváme speciální morfologickou značku, která je od zkratk *Gsg* a *Gpl* figurujících ve výčtu pádových forem odlišena pomocí dolního indexu.

**bezmaksas** <sub>Gsg</sub> *bezplatný* [...]  
**bezsapņu** <sub>Gpl</sub> *bezesný* [...]  
**dziļūdens** <sub>Gsg</sub> *hlubinný* [...]  
**daudzbērnu** <sub>Gpl</sub>: • **d. ģimene** *rodina s mnoha dětmi, početná rodina*

Na rozdíl od některých překladových slovníků genitivní tvary lexémů majících kompletní deklinační paradigma nevyčleňujeme jako samostatná lemmata, byť by byly frekventované, ale zařazujeme je do běžných kolokací pod příslušné lemma.

<sup>167</sup> Vyčleňovány jsou též v lotyšském tagsetu. Pomocí dotazu [tag="N...s."] získáme v LVK 3534 výsledků sdružených do 319 lemmat, z nichž dvacítku nejčastějších jsou tyto formy: *vācu* 376, *ārkārtas* 302, *krājizdevu* 226, *pirmsskolas* 176, *dzimtsarakstu* 133, *bezmaksas* 129, *daudzdzīvokļu* 116, *plašsaziņas* 98, *pašmāju* 96, *brīvdabas* 57, *starpvalstu* 57, *priekšvēlēšanu* 55, *pirmstiesas* 55, *daudzstāvu* 50, *pārrobežu* 46, *latīņu* 46, *piemājas* 40, *daudzbērnu* 38, *rakstveida* 37, *pretepidēmijas* 36.

<sup>168</sup> Ojediněle též jako adverbiale: *bezgala plašs* 'nekonečně / nesmírně široký' – *bezgala plašums* 'nekonečná šíře'.

**ārzelmes** f<sub>5</sub> pl (Gpl ~mju) *zahraničī, cizina* [...] • ~mju ministrs *ministr zahraničī* • ~mju valūta *cizī valuta* • ~mju ceļojums *cesta do zahraničī (ciziny)* [...]

**bezmiegls** m<sub>1</sub> *nespavost* [...] • ~a nakts *bezesnā (probdēlā) noc*

Početnou skupinu dubletních složenin, z nichž jedna je ohebným adjektivem, druhá nesklonným tvarem Gsg a které jsou většinou chápány jako plně synonymní, zpracováváme konzistentně v rámci stejného lemmatu; na prvním místě je uvedena frekventovanější varianta, v případě neprůkaznosti korpusových dat neohebná podoba:

**bezalkoholisk[s]** [o-o:] adj; **bezalkohola** [o-o:] Gsg (nápoj) *nealkoholický*

**konusveida** [o] Gsg; **konusveidīg[s]** [o] adj *kuželovitý, kuželový, kónický*

V dnešní lotyštině je třída ustrnulých genitivních forem velmi produktivní, obzvláště v oblasti terminologie, publicistiky a beletrie. Souvisí to s typickým rysem lotyštiny používat genitivní tvary tam, kde se v jiných jazycích včetně češtiny uplatňují relační adjektiva (LVG: 369). Tuto odlišnost je nutno vystihnout vhodnou exemplifikací. Méně časté tvary a okazionalismy (v kombinaci s číslovkami teoreticky neomezené množiny slov) se LČ podchytit nesnaží, v takových případech je uživatel nucen správně rozdělit danou složeninu (např. *četršķautņi* „čtyřhranný, se čtyřmi hranami“ s jediným výskytem v LVTK) na jednotlivé části a ty pak hledat pod příslušnými hesly (*četri*, resp. *četr–*, *čtyř–* a *šķautnis*).

### 3.2.3.2 Gramatická charakteristika verb v LČ

Z gramatického systému lotyšského verba věnujeme pro účely této kapitoly pozornost pouze vybraným jevům, zajímavým z hlediska kontrastivního, tj. těm, jež se v lotyštině a češtině vzájemně liší a při lexikografickém zpracování lotyšských sloves v LČ mohou působit problémy. Asymetričnost verbálních sys-

témů v obou jazycích je nutno kompenzovat vhodnými prostředky, které se pokoušíme najít. V následujících odstavcích se zaměříme na kategorii času a s ním souvisejícího vidu, dále se budeme věnovat slovesnému způsobu a systému konjugačních tříd a konečně bohatému rejstříku participií.

Lotyšská verba jsou v hesláři LČ reprezentována jednak infinitivním tvarem, končícím na *-t* (u zvratných sloves na *-ties*), jednak jsou jako odkazová hesla zpracovávány kromě supletivních forem soustavně též přítomní a přeteritální kmeny, v nichž dochází oproti kmenu infinitivnímu k hláskové alternaci. Tuto praxi přebíráme od A. V. Andronova (2002), lotyšské překladové slovníky alternované slovesné kmeny do hesláře vesměs nezahrnují, s výjimkou vybraných supletivních tvarů, a uživatel je tak odkázán na případný soupis „obtížnějších“ (ČLN: 411) sloves vně hesláře. Že přesná identifikace lemmatu, již je správný překlad lexému podmíněn, není triviální, ukazuje názorně následující příklad:

**liec<sup>1</sup>, liecies** préz ↓likt<sup>1</sup>, likties<sup>1</sup>, likt<sup>2</sup>, likties<sup>2</sup>  
**liec<sup>2</sup>, liec-** préz ↓liekt, liekties  
 ...  
**liek, liek-** préz ↓likt<sup>1</sup>, likties<sup>1</sup>, likt<sup>2</sup>, likties<sup>2</sup>  
 ...  
**lie|kt** v<sub>1</sub> (~cu, ~c, ~c; ~cu) **1.** (ko) *ohýbat* × *ohnout* (co) [...]  
**lie|kties** v<sub>1</sub> (~cos, ~cies, ~cas; ~cos) **1.** *ohýbat se* × *ohnout se*, *sklánět se* × *sklonit se* [...]  
 ...  
**likt<sup>1</sup>** v<sub>1</sub> (lieku, ~c, ~c; ~cu) **1.** (ko kur) *pokládat*, *klást* (co kam) [...]  
**likt<sup>2</sup>** v<sub>1</sub> (lieku, ~c, ~c; ~cu) **1.** (ko kam) *nařizovat* × *nařídít*, *přikazovat* × *přikázat* (co komu) [...]  
**līk|t<sup>1</sup>** v<sub>1</sub> (~stu, ~sti, ~st; ~a) **1.** *ohýbat se* × *ohnout se*, *sklánět se* × *sklonit se* [...]  
 ...  
**likties<sup>1</sup>** v<sub>1</sub> (liekos, ~cies, ~kas; ~cos) **1.** *lehat si* × *lehnout si*, *ulehat* × *ulehnout* [...]  
**likties<sup>2</sup>** v<sub>1</sub> (liekos, ~cies, ~kas; ~cos) **1.** (kam ko) *zdát se*, *jevit se* (komu co) [...]

Jak patrně, u slovesných kmenů je uveden jejich typ (přítomný, či minulý; infinitivní je zastoupen samotným infinitivem) a uživatel je odkázán na odpoví-

dávající infinitiv. Ten je přiřazen k jedné ze tří konjugačních tříd a následně je podán výčet morfologických tvarů, na jejichž základě lze odvodit celé paradigma: 1. osoba singuláru indikativu přítomného času, 2. osoba singuláru indikativu přítomného času, 3. osoba singuláru/plurálu<sup>169</sup> indikativu přítomného času a 1. osoba singuláru indikativu minulého času; u sloves neosobních pak pouhé dva tvary: 3. osoba indikativu přítomného času následovaná 3. osobou indikativu minulého času.<sup>170</sup> V případě nepravidelného tvoření zařazujeme další tvary (futurum, participia aj.) s příslušnou zkratkou; podány jsou rovněž dubletní formy. Kompletní paradigmata pro všechny konjugační třídy (o nich dále v 3.2.3.2.4) budou uvedena v přehledu lotyšské gramatiky na konci slovníku.

### 3.2.3.2.1 Slovesný čas

Systém slovesného času v lotyštině je oproti češtině složitější: vedle trojice jednoduchých časů (přítomný, minulý, budoucí), jejichž použití je v obou jazycích až na dílčí výjimky stejné, zahrnuje ještě tři časy složené (perfektum, pluskvamperfektum, budoucí II), které mohou českému uživateli dělat potíže. Tvoří se kombinací příslušného tvaru pomocného slovesa *būt* s participiem minulého času aktiva a vyjadřují primárně stav vzniklý v důsledku předchozího děje v přítomnosti, minulosti či budoucnosti, a to nejčastěji ve vztahu k jinému ději či

<sup>169</sup> Lotyšské sloveso rozlišuje tři osoby v singuláru a tři osoby v plurálu, přičemž *syntetické* tvary 3. osoby obou čísel jsou vždy shodné, číslo je dáno podmětem věty. Restriktivní přívlastek, zdůrazněný v předchozí větě kurzívou, je tu důležitý (a LVG: 516 jej správně zmiňuje); obecné tvrzení, že lotyšský slovesný systém má společný tvar pro 3. osobu, tradované v české balistice (ČLS, Erhart 1984: 125; též o litevštině 100), je totiž nesprávné, jak lze snadno ukázat na následujících větách, kde se tvary 3. osoby (proložené tučně) liší nejen v čísle, ale též v rodě:

*Žurnāls tiek lasīts.* „Román je čten / se čte.“ – *Grāmata tiek lasīta.* „Kniha je čtena / se čte.“

*Žurnāli tiek lasīti.* „Romány jsou čteny / se čtou.“ – *Grāmatas tiek lasītas.* „Knihy jsou čteny / se čtou.“

*Viņš ir šo grāmatu lasījis.* „Četl tu knihu.“ – *Viņa ir šo grāmatu lasījis.* „Četla tu knihu.“

*Viņi ir šo grāmatu lasījuši.* „Četli tu knihu.“ – *Viņas ir šo grāmatu lasījušas.* „Četly tu knihu.“

<sup>170</sup> Stejně tvary uvádějí i hlavní lotyšské výkladové slovníky, srov. též pět (tři u neosobních sloves) forem uváděných v LP i LL: první tři jsou shodné s těmi uváděnými v LČ, präteritální tvar je zastoupen 3. (nikoliv 1.) osobou, mimoto je uváděno také futurum (opět tvar 3. osoby), které se však tvoří pravidelně, a až na určité výjimky, na něž LČ výslovně upozorní, nepředstavuje problém. Naproti tomu A. Andronov (2002) upřednostňuje ve výčtu slovesných tvarů 3. osobu, resp. vystačí si s pouhými dvěma jejími tvary: přítomným a präteritálním (na rozdíl od LČ pak ale musí explicitně zkratkou vydělovat neosobní slovesa). Jeho návrh je dosti problematický, sám autor oproti dvěma nesporným případům, kdy je 3. osoba nejreprezentativnější z celého paradigmatu, přiznává hned tři slabiny (Andronovs 1998: 147–148).

stavu. Jedná se o časy relativní, na rozdíl od absolutních – jednoduchých, jež se vztahují přímo k okamžiku promluvy.<sup>171</sup>

Složený čas přítomný (perfektum) slouží k vyjádření děje již dokončeného, jehož výsledek trvá do okamžiku promluvy.

Složený čas minulý (pluskvamperfektum) označuje děj či jeho výsledek, který proběhl nebo nastal před jiným dějem (stavem) v minulosti; může odkazovat rovněž k ději proběhlému ve vzdálenější minulosti, jehož důsledky jsou pro přítomný okamžik irelevantní.

Složený čas budoucí (futurum II) vyjadřuje výsledek děje či stav, který proběhne nebo nastane před jiným budoucím dějem (stavem).

Složené časy bývají v běžném jazyce často nahrazovány odpovídajícím časem jednoduchým;<sup>172</sup> v přehledu lotyšské gramatiky, jež má být součástí LČ, musí být jejich užití popsáno včetně vhodných ilustrativních příkladů s českými překlady, již vzhledem k tomu, že možnosti naznačit jejich sémantiku, zvláště významy okrajové, na úrovni mikrostruktury slovníku jsou výrazně omezené.

#### 3.2.3.2.2 Slovesný vid

Kategorie vidu je nejproblematictější ze všech gramatických kategorií lotyšského slovesa: LVG opakovaně postuluje její spornost, odvíjející se od nedostatečné míry formalizace (LVG: 531) a na slovesný vid nahlíží „v souvislosti s různými [...] vyjadřovacími prostředky – gramatickými, derivačními a morfonologickými –, upouštějíc od popisu vidu jakožto gramatické kategorie“ (ibid.), a to na základě nedostatečné abstrakce gramatických znaků,

---

<sup>171</sup> Ve složených časech se nicméně může vyskytnout i imperfektivní tvar slovesa, v takových případech se dostává do popředí průběh děje, důraz je kladen na jeho délku, trvání či opakování.

<sup>172</sup> Občasné nahrazování perfekta préteritem lze připsat na vrub odchylkám jednotlivých mluvčích od normy, masověji se však, i přes původní obavy některých lingvistů (Lokmane [1988: 119] zaznamenává, že pod vlivem ruštiny se v lotyštině začínají rozmáhat jednoduché slovesné časy na úkor složených, a její článek vyznívá varovně: „[...] přijde-li lotyšтина o složené časy, stane se chudší.“) nerozšířilo; v časech relativních je pak paralelismus jednoduchých a složených časů dokonce součástí spisovného jazyka.

Naopak, vzájemná konkurence obou forem daná tím, že každá akcentuje něco jiného – jednoduché časy průběhovou, kdežto složené rezulativnost –, se jeví pro jazyk jako dostatečně nosná, složené formy tak zůstávají charakteristickým gramatickým rysem lotyštiny a jejich použití je ve většině kontextů nepříznakové.



závislosti vidu na lexikálním významu slovesa (sémantika prefixů, v menší míře též sufixů) a také jeho finální realizace často až v daném kontextu (ibid.: 532).<sup>173</sup> Pro nás je však podstatné, že v lotyštině existuje možnost charakterizovat děj z hlediska jeho ne/dokončenosti (ne/dovršenosti, ne/uzavřenosti) a že tu existuje nezpochybnitelná systémová korelace (MLLVG I: 567; LVG: 533; Kalme–Smiltniece 2001: 219). Budeme proto nadále, s vědomím výše naznačených specifik, hovořit právě o vidu a v návaznosti na to i o slovesech perfektivních a imperfektivních.

LVG řadí im/perfektiva spolu se slovesy ne/přechodnými, zvratnými a kauzativními do lexikálně-gramatických skupin, úzce svázaných s jinými slovesnými kategoriemi, především časem a způsobem. V lotyšských výkladových slovnících pak není zvykem vid označovat,<sup>174</sup> z pohledu rodilého mluvčího jde o informaci nadbytečnou, neboť „prefigovaný a neprefigovaný tvar slovesa s týmž kořenem je prvním vnějším znakem perfektivnosti a neperfektivnosti“ (MLLVG I: 565); neprefigovaná slovesa jsou vnímána jako imperfektivní, kdežto slovesa prefigovaná jako perfektivní.

I kdybychom pominuli řadu výjimek z tohoto základního principu,<sup>175</sup> je nezbytné při zpracování LČ toto východisko korigovat s ohledem na perspektivu českého uživatele, pro nějž je vidová distinkce mnohem ostřejší. Lotyšským neprefigovaným slovesům odpovídá v češtině prakticky pokaždé imperfektivní slovesný tvar; úkolem lexikografa je tedy najít *sémanticky* co nejpřesnější ekvivalent. U sloves prefigovaných, splývajících v lotyštině povětšinou s perfektivností, je situace složitější: vedle perfektivního slovesného tvaru je můžeme téměř vždy překládat příslušným sekundárně vytvořeným imperfektivem (*braukt* ‚jet‘, též ‚jezdit‘, kdežto *atbraukt* ‚přijet × přijíždět‘, srov. LR: *examь*;

<sup>173</sup> Podobná situace panuje i v litevské jazykovědě (ibid.: 457, Mathiassen 1996).

<sup>174</sup> Výjimkou je pouze 1. svazek LLVV, počínaje 2. svazkem bylo však od této praxe ustoupeno.

<sup>175</sup> Kromě vidových dvojic existují také systémové nepravidelnosti, známé i v češtině: slovesa obouvidová (zvláště internacionalismy: *publicēt* ‚publikovat‘, ale i slova domácího původu: *pārdot* ‚prodávat × prodat‘), perfektiva tantum (např. sloveso *iedegties* nemá svůj imperfektivní protějšek, byť ho do češtiny můžeme přeložit jako ‚rozhořet se‘, ale i ‚rozhořívát se‘) i imperfektiva tantum (*atrasties* ‚nacházet se‘) a též sekundární imperfektiva (*braukāt* ‚jezdit, jezdívát‘)

*npueхать-npueзжать*; LP: *jechać; przyjechać / przyjeżdżać*), a to dokonce i v těch případech popisovaných v gramatikách, kdy má mít prefix čistě perfektivizační funkci.<sup>176</sup> V této hypotéze nás utvrzují oba námi užívané překladové slovníky jazyků, které slovesný vid rozlišují (LP, LR) – ty většinu lotyšských perfektiv svorně překládají jak dokonavým protějškem, tak jeho sekundárním nedokonavým korelátem –, ale svým způsobem rovněž oba česko-lotyšské slovníky (ČL, ČLN), hnízdující běžně české vidové dvojice pod jedno heslo, a ostatně svědčí o tom i naše vlastní zkušenost s dosud zpracovanými hesly.

Ideální postup spatřujeme ve zpracovávání prefigovaných sloves spolu s jejich neprefigovaným protějškem, což může mimo jiné ukázat na odlišnou kolokabilitu a z ní plynoucí sémantiku, neboť vidová souvztažnost se může odlišovat od té věcně významové.<sup>177</sup> Zároveň je účelné zpracovávat systematicky slovesa se stejnými prefixy, snažit se následně najít sémantiku jednotlivých prefixů a zkoumat, nakolik odpovídá sémantice paralelních českých předpon.<sup>178</sup>

České vidové protějšky budou v LČ odděleny speciálním znakem pro násobením  $\times$ ,<sup>179</sup> vyhrazeným pouze pro tento případ, aby nedocházelo k nežádoucí polysémii lexikografických značek, a charakterizovány příslušnou zkratkou (*i* pro imperfektiva, *p* pro perfektiva, *i/p* pro slovesa obouvidová) v dolním indexu: jde o informaci určenou primárně pro lotyšské uživatele.

**abonēt**[t] [o] v<sub>2</sub> (~ju, ~, ~; ~ju) (ko) (tiskoviny, divadlo, služby ap.) *předplácet* [si]<sub>i</sub>  $\times$  *předplatit* [si]<sub>p</sub>,  
*abonovat* <sub>i/p</sub> (co) || (tiskoviny též) *odebírat*<sub>i</sub> (co) || (služby též) *objednávat* [si]<sub>i</sub>  $\times$  *objednat* [si]<sub>p</sub>  
 (co) || (k odběru) *přihlásit se*<sub>p</sub> (k čemu/kam) • a. avízi / žurnálu předplatit si noviny / časopis

<sup>176</sup> Je však třeba konstatovat, že tyto případy jsou dost řídké, naopak dominuje sémanticko-slovotvorná funkce prefixace, kdy „perfektivizace se uskutečňuje jen jako »vedlejší produkt« slovotvorného procesu“ (Lokmane 1998, 111), předpona základní význam slovesa modifikuje (z hlediska časového, prostorového či kvantitativního), čímž vzniká nový lexém, mající v češtině nejčastěji oba vidové ekvivalenty. Blíže ke dvojí funkci prefixu viz Škrabal 2011: 264.

<sup>177</sup> Viz např. odlišnou sémantiku slovesného páru *dostat*  $\times$  *dostávat* v ČLN (57): první forma má 8 významů, druhá pouze 3; zvrtná podoba *dostat se* rovněž 8 významů, kdežto *dostávat se* je zastoupeno pouze kolokací *nedostává se mi slov*.

<sup>178</sup> Prefixy by měly tvořit v hesláři LČ samostatná hesla s uvedením perfektivizační funkce jako jedním z významů (viz též 3.1.2.4.2).

<sup>179</sup> Srov. užití tiretu (-) v LR a lomítka (/) v LP.

V překladech kolokací už pak postačí uvádět jedinou, bezpříznakovou formu, přičemž druhá zůstává vyjádřena implicitně: *atbraukt mājās* ‚přijet domů‘ (implicitě též ‚přijíždět‘); *pārdot preces* ‚prodat zboží‘ (implicitě též ‚prodávat‘).<sup>180</sup> Postihnout vidovou distinkci na vyšší než kolokační rovině, kde se však obvykle projevuje, je pro lexikografa, snažícího se zachytit především ustálenou, systémovou podobu jazyka, úkol obtížný, ne-li nemožný (podrobněji o možnostech a mezích slovníku v tomto ohledu Škrabal 2011, 2012). Gramatikami uváděné<sup>181</sup> příkladové věty, které tuto distinkci reflektují, nelze z čistě praktických důvodů do slovníku implementovat; v žádném nám známém překladovém slovníku podobné věty také nefigurují.

3.2.3.2.6.1 Pod kategorii vidu se v lotyšské jazykovědě řadí také opozice mezi jednorázovými a vícerázovými slovesy (*vienkārtējas / daudzkārtējas darbības veids*), případně též dalšími spadajícími pod způsob slovesného děje (*Aktionsart*). Iterativa se od semelfaktiv odvozují morfonologicky: pomocí sufixu (*braukt : braukāt*) či interfixu (*liet : liedēt*) nebo vokalickou (*vilkt : valkāt*) či intonační změnou (*slāukt : slaūcīt*), případně kombinací obou těchto principů (*smiēt : smaīdīt*; všechny příklady z LVG: 540). Na jejich vazbu není v LČ třeba explicitně poukazovat (podobně jako např. vazbu kauzativ a faktitiv), cílem je najít co nejpřesnější český ekvivalent, jen v případě jejich totožnosti je opatřit vysvětlující glosou, aby českých uživatel danou dvojici lotyšských lexémů nepovažoval mylně za synonyma. Navíc se iterativa mohou dále pojit také s prefixy, byť sémantika opakovaně vykonávané činnosti se tu nutně střetává s perfektivností.

**stu|mt** v<sub>1</sub> (~mju, ~m, ~m; stūmu) (*ko [kurp]*) *tlačit<sub>i</sub>* × *zatlačit<sub>p</sub>*, *strkat<sub>i</sub>* × *strčit<sub>p</sub>*, *sunout<sub>i</sub>* × *posunout<sub>p</sub>* (*co/koho [kam]*) [...]

**stumd|īt** v<sub>3</sub> (~u, ~i, ~a; ~īju) (*ko kurp*) *postrkovat<sub>i</sub>* × *postrčit<sub>p</sub>*, *posunovat<sub>i</sub>* × *posunout<sub>p</sub>* (*co/koho kam*) [...]

<sup>180</sup> Srov. LR *atbraukt* при́ехать-прие́зжат; a. *ciemos* при́ехать в гости; LP (171) *atbraukt* przyjechać / przyjeżdżać; a. *mājās* przyjechać (wrócić) do domu

<sup>181</sup> Viz též jiné příklady, založené na datech z česko-lotyšské části IC – Škrabal 2011: 264–267

<sup>?</sup>*pastumd|īt* v<sub>3</sub> (~u, ~i, ~a; ~īju) (ko *kurp*) <sup>?</sup>*popostrčit<sub>p</sub>* × <sup>?</sup>*popostrkávati*, <sup>?</sup>*poposunout<sub>p</sub>* × <sup>?</sup>*poposunovat<sub>i</sub>* (co/koho kam) [...] <sup>182</sup>

3.2.3.2.6.2 Vidová problematika se však netýká pouze sloves. Kromě logicky četných deverbativních derivátů podstatně zasahuje též adverbia. Roli vidového protějšku sloves s prefixem lokálního významu může mít daný neprefigovaný tvar s postponovaným adverbium, popřípadě též některými substantivy. <sup>183</sup> Tak se u výše zmíněného verba *stumt* kromě kolokátů v roli předmětu (*stumt galdu, skapi, ratiņus...* ‚stůl, skříň, vozík...‘) často objevují adverbia lokálního významu, které mohou dané sloveso perfektivizovat a konkurují tak některému z prefixů.

*stumt ārā* ‚vystrčit, vytlačit × vystrkovat, vytlačovat‘, dosl. ‚tlačit ven‘ ~ *izstumt*

*stumt iekšā* ‚vstrčit, vtlačit × vstrkovat, vtlačovat‘, dosl. ‚tlačit ven‘ ~ *iestumt*

*stumt klāt* ‚přistrčit, přitlačit × přistrkovat, přitlačovat‘, dosl. ‚tlačit k‘ ~ *piestumt* aj.

Slovesný vid je tu vyjádřen syntakticky, kontextuálně (což je nejlépe vidět v rámci celé věty): syntaktické imperfektivum stojí v opozici k morfologickému, prefigovanému perfektivu (LVG: 536). Přestože jde o periferní způsob vyjadřování slovesného vidu v lotyštině (ibid.: 532, 537, srov. Kalnača 2004), měl by být v LČ náležitě reflektován; typické je především pro běžně mluvenou lotyštinu, není však třeba označovat toto užití klasifikátorem.

**stul|mt** v<sub>1</sub> (~mju, ~m, ~m; stūmu) (ko [*kurp*]) *tlačit<sub>i</sub>* × *zatlačit<sub>p</sub>*, *strkat<sub>i</sub>* × *strčit<sub>p</sub>*, *sunout<sub>i</sub>* × *posunout<sub>p</sub>* (co/koho [*kam*]) • s. *ratiņus tlačit vozík* • s. *durvis tlačit do dveří* • s. *rokas kabatā strčit [sij] ruce do kapsy* || • s. *ārā*, s. *projām* (prom) *vystřčit<sub>p</sub>* × *vystrkovat<sub>i</sub>*, *vytlačit<sub>p</sub>* × *vytlačovat<sub>i</sub>* • s. *iekšā vstrčit<sub>p</sub>*, *vtlačit<sub>p</sub>* • s. *klāt pīstrčit<sub>p</sub>* × *pīstrkovat<sub>i</sub>*, *pīttlačit<sub>p</sub>* × *pīttlačovat<sub>i</sub>*

<sup>182</sup> V takovýchto případech už ale nutně budeme narážet na nedostatek korpusových dat, vedoucích k elementární otázce, zda daný lexém vůbec zařadit do hesláře. Viz LVK: lemma *stumt* 61 výskytů, *stumdīt* 8, *pastumdīt* 1; LVTK *stumd.* \* 112, *pastumd.* \* 9. Srov. např. čtyřčlennou řadu *braukt* ‚jet, jezdit‘ – *braukāt* ‚jezdit, jezdivat‘ – *braukalēt* ‚jezdit, jezdivat [obv. bez vážného důvodu]‘ – *pabraukalēt* ‚popojezdit si‘, z níž do překladových slovníků bývají běžně zahrnuty jen první dvě formy, srov. též v LVTK *braukal.* \* 21, *pabraukal.* \* 5, v LVK je zastoupeno 1 výskytem jen deverbativum *braukalēšana*, k němuž bychom sotva našli český ekvivalent (ekvivalence by se tak musela hledat na vyšší rovině).

<sup>183</sup> Jde vesměs o tematický okruh oblékání, substantiva jsou omezena na názvy částí těla a vyskytují se v dané konstrukci pouze v lokativu, např. *likt cepuri galvā* ‚nasadit si čepici [na hlavu]‘, *vilkt mēteļi mugurā* ‚obléct si kabát‘, dosl. ‚navléct si kabát na záda‘)

Dvojitá svislice vizuálně vhodně odděluje obě skupiny kolokátů, nejde přitom o samostatný význam, jež by bylo nutno uvodit arabskou číslicí. Naproti tomu u samotných adverbíí by měl tento význam oddělen jako suverénní.

**ārā** adv **1.** (kam) *ven* [...] || **2.** (kde) *venku* [...] || **3.** (ve spojení s nepředponovými slovesy odpovídá předponě *iz-*) • stumt ā. (≡ *izstumt*) *vystřčit<sub>p</sub>* × *vystrkovat<sub>i</sub>*, *vytlačit<sub>p</sub>* × *vytlačovat<sub>i</sub>*[...]

### 3.2.3.2.3 Slovesný způsob

Lotyština rozlišuje celkem pět způsobů: kromě trojice shodných s češtinou (indikativ, imperativ, kondicionál) navíc dva specifické: debitiv (způsob „nutnostní“: *vajadzības izteiskme*) a relativ (způsob „vyprávěcí“: *atstāstījuma izteiksmē*).<sup>184</sup> Všechny budou i s příklady popsány v gramatickém přehledu, zde zmiňujeme pouze poslední dva:

3.2.3.2.3.1 Relativ slouží nejčastěji k reprodukci výroku jiné osoby, zvláště není-li si mluvčí jist jeho pravdivostí, případně chce-li se od něj distancovat (funkčně se tak podobá např. německému konjunktivu II); do češtiny jej lze přeložit slovem *prý*, *asi* apod., případně opisně. Pojí se se čtyřmi časy: jednoduchým přítomem a futurem a složeným přítomem a futurem. V prvních dvou případech jde o neohebný tvar tvořený od příslušného času indikativu aktiva nahrazením osobních koncovek neměnným sufixem *-ot* (*-oties* u reflexiv), v případě složených časů kombinací tvaru *esot* (= relativ jednoduchého přítomného slovesa *būt*) či *būšot* (= relativ jednoduchého futura slovesa *būt*) s přičestím minulým činným plnovýznamového slovesa. Naopak lotyšský relativ nerozlišuje osobu a číslo, ty jsou odvoditelné pouze z kontextu.

3.2.3.2.3.2 Debitivem se v lotyštině vyjadřuje nutnost nebo potřeba daný děj uskutečnit; do češtiny se převádí nejčastěji způsobovým slovesem *mušet*, eventuálně jiným, obdobným prostředkem. Kombinuje se se všemi časy, oběma

<sup>184</sup> Takto popisují kategorii způsobu poslední dvě lotyšské gramatiky (MLLVG I 1959; LVG 2015), stejně ji pojmají i české baltistické prameny (Erhart 1984, ČLN), s výjimkou Jiřího Marvana (Marvans 1967), zpochybňujícího samostatný charakter debitivu.

rody, ze způsobů s relativem a kondicionálem, (ne všechny tvary jsou však stejně užívané). Tvoří se spojením takové kombinaci odpovídajícího tvaru 3. osoby pomocného slovesa *būt* včetně negativních forem (přezentní tvar *ir* se však často vypouští) se zvláštním debitivním tvarem: 3. osoby přezentu plnovýznamového slovesa s prefixem *jā*–.<sup>185</sup> Původce debitivního děje se vyjadřuje dativem, předmět děje nominativem (až na osobní rodová zájmena *es*, *tu*, *mēs*, *jūs* a zvrátne *sevis*, které mají v této pozici akuzativní tvar<sup>186</sup>).

K debitivu bývají v lotyšských gramatikách tradičně řazeny i jeho sémantické varianty: konstrukce modálních sloves *vajadzēt*, *nākties* či pomocného slovesa *būt* s infinitivem a užití participia přezentu pasiva ve slovesně-jmenném přísudku. Tuto funkci daných sloves lze naznačit přímo v hesle, samostatně je zpracován i prefix *jā*–:<sup>187</sup>

**vaja|dzēt** v<sub>3</sub> (~g; ~dzēja) vyjadřuje nutnost, nezbytnost, potřebu ap.<sup>188</sup> • man ~g naudas *potřebuju peníze* • ko tev ~g? *co potřebuješ?* • ko jums ~g? (v obchodě, na úřadě ap.) *co byste potřeboval?* • viņai drīz ~g būt atpakaļ *měla by se už brzo vrátit* • mums ~dzētu satikties *měli bychom se sejít* • valdībai ~dzēs rēķināties ar.. *vláda bude muset počítat s...* • tev ~g / ~dzēs būt uzmanīgākam *musíš / budeš muset být opatrnější [...]* || *být potřeba, být zapotřebí* • ~g

<sup>185</sup> S výjimkou samotného slovesa *būt*, jehož debitiv je *jābūt*.

<sup>186</sup> Akuzativu namísto nominativu se používá běžně i v běžně mluvené řeči, zvláště je-li součástí předmětu infinitiv (LVG: 487).

<sup>187</sup> Debitiv je tak jediný slovesný způsob, který lze v LČ postihnout na makrostrukturní úrovni.

<sup>188</sup> Řešení nabízená ostatními slovníky jsou sporná. Zdánlivá ekvivalence mezi infinitivním tvarem *vajadzēt* a třemi protějšky nabízenými LP (*potrzebować*; *musieć*; *mieć powinność*) je v přímém rozporu s korpusovými daty. Viz statistika tvarů v LVK, v němž je lemma *vajadzēt* doloženo celkem 2578 výskytů, z toho:

*vajag* (přezens) 904

*vajadzētu* (kondicionál) 785

*vajadzēja* (préteritum) 625

*vajadzēs* (futurum) 170

*vajadzējis* (participium préterita aktiva) 52

*vajagot* (relativ) 38

*vajadzēt* (infinitiv) 2

*vajadzēšot* (relativ futura) 2

Infinitiv se tedy prakticky nevyskytuje (realizovat se může pouze ve spojení s jiným modálním slovesem či metajazykově), a těžko tedy k němu hledat ekvivalent na úrovni lemmatické. Ekvivalentace úspěšně proběhne až na úrovni kolokační: např. *man / tev / mums vajag* ‚potřebuju / potřebuješ / potřebujeme‘ (dosl. ‚mně / tobě / nám potřeba‘). Je tedy patrně lepší rezignovat na přímý ekvivalent a nahradit jej popisem funkce, ostatně jde o modální sloveso, kde sémantika ustupuje právě funkci, mimoto sém nutnosti/potřeby tu zůstane zachován.

Další možné řešení je opustit infinitiv jako reprezentativní tvar a za lemma prohlásit tvar nejfrekventovanější, tedy *vajag* (srov. obdobný případ s problematickým lemmatem v kap. 4.3). Ani tak nepokročíme příliš dál, také tento tvar – a s ním i všechny formy paradigmatu – se plně realizuje až na úrovni kolokační, navíc přibude problém s odvozováním jiných alolexů (od tvaru *vajag* to není možné).

Optimální řešení vidíme v glose popisující funkci namísto v umělém ekvivalentu, a je-li k dispozici takto průkazná statistika, také v jejím uvedení na konec heslové stati (viz 3.2.7).

pieņemt lēmumu *je potřeba přijmout rozhodnutí* • visu ~g padarīt laicīgi *všechno je potřeba udělat včas* [...] || ♦ tā tev ~g (~dzēja) *dobře ti tak, to ti patří*

**jā–** prefix debitivu, ve spojení s 3. osobou přítomného času vyjadřuje nutnost, nezbytnost, potřebu ap., např. *man [ir] jāiet* („musím jít“), *tev ir mani jāapciemo* („musíš mě navštívit“), *viņai būs jā atrod citu vietu* („bude muset najít jiné místo“) → **Gramatika: s. 1052**

3.2.3.2.4 Slovesa v lotyštině se na základě počtu slabik v jednotlivých slovesných kmenech dělí do tří konjugačních tříd, jejichž přehled včetně tvaroslovných paradigmat bude zahrnut do gramatické přílohy LČ. Podrobnější klasifikaci, prezentovanou nejnovější lotyšskou gramatikou (LVG: 545–549),<sup>189</sup> nepřejímáme, pro českého uživatele je dostačující informace o příslušné konjugaci, ta spolu se základními morfologickými tvary, jimiž bude vybaveno každé neodkazové slovesné lemma, dovoluje rekonstruovat celé paradigma; informace o příslušnosti k té či oné skupině je nadbytečná, mimoto se může v závislosti na zdroji a době lišit.

Takto vydělujeme slovesa první konjugace, u nichž infinitivní koncovka *–t(ies)* navazuje přímo na slovesný kořen, od sloves druhé a třetí konjugace, majícími mezi kořenem a infinitivní koncovkou kmenotvorná přípona. Slovesa druhé a třetí konjugace se od sebe liší tvořením tvarů přítomného času (Kalnača 2013: 75).

3.2.3.2.4.1 Do první konjugace se řadí slovesa primární, která v infinitivním kmenu nemají sufix (s výjimkou zvratného formantu *–ties*). Přítomný a minulý kmen se může s kmenem infinitivním shodovat, anebo se od něj odvozuje různými prostředky: nejčastěji hláskovou změnou či přidáním sufixu (*–j–*, *–n–* nebo *–st–*); v případě, že se kmen přítomný či minulý odlišuje od kmenem infinitivního (tedy i od lemmatu uvedeného v hesláři) a jeho správné dekódo-

<sup>189</sup> Zde je 1. konjugace dělena do tří skupin, 2. konjugace se dále nečlení, 3. konjugace má dvě podskupiny. Tento systém je převzat od E. Soidové a přejímá jej i většina gramatických popisů lotyštiny z poslední doby (Kalme–Smiltnece 2001: 211–216; Paegle 2003: 104–111; Kalnača 2013: 75–78), nahradil tak dřívější popis v MLLVG (5 tříd v 1. konjugaci, 4 třídy v druhé, 2 třídy ve třetí). Kombinací obou je pojetí nabízené v ČLN (5 tříd v 1. konjugaci, 2. konjugace dále nečleněna, 2 třídy ve 3. konjugaci), srov. též Erhartovo pojetí (Erhart 1984: 126–128). Jinak se však v heslářích většiny překladových slovníků otázka konjugací neřeší (LR, LL, Andronov 2002), a pokud ano, pak bez dalšího členění na třídy (LP).

vání by uživateli mohlo činit obtíže, uvádíme jej jako samostatné, odkazové heslo (viz výše příklad *lik-* / *liek-*). Slovesa s infinitivem zakončeným na *-st(ies)* či *-zt(ies)* tvoří budoucí čas od kmene minulého pomocí infixu *-ī-*, takovéto formy zahrnujeme do výčtu morfologických tvarů.

**nest** *v*<sub>1</sub> (nesu [æ], nes, nes [æ]; nesu; fut nesīšu) **1.** (*ko*) *nést, nosīt (co)* [...]

**vest** *v*<sub>1</sub> (vedu [æ], ved, ved [æ]; vedu; fut vedīšu; partic vezdams) **1.** (*ko [kurp]*) *vést, vodīt (koho [kam])* [...]

**grie|zties** *v*<sub>1</sub> (~žos, ~zies, ~žas; ~zos; fut ~zīšos) **1.** *točit se, otáčēt se* × *otočit se, obracet se* × *obrátīt se* [...]

3.2.3.2.4.2 Slovesa druhé konjugace jsou ve všech třech kmenech tvořena sufixem *-ā-*, *-ē-*, *-ī-* či *-o-*,<sup>190</sup> v kmenu přítomném i préteritálním navíc sufixem *-j-*. V přítomnu se používá redukovaná série osobních koncovek: 2. osoba singuláru a 3. osoba obou čísel ztrácí přítomný sufix i osobní koncovku.

**runā|t** *v*<sub>2</sub> (~ju, ~, ~; ~ju) **1.** *mluvīt, hovořīt* [...]

**dzīvo|t** *v*<sub>2</sub> (~ju, ~, ~; ~ju) **1.** *žīt* [...]

**abonē|t** [o] *v*<sub>2</sub> (~ju, ~, ~; ~ju) (*ko*) (tiskoviny, divadlo, služby ap.) *předplācēt [si]* × *předplatīt [si]*, *abonovat (co)* [...]

3.2.3.2.4.3 Slovesa třetí konjugace mají infinitivní kmen rozšířený o sufix *-ī-*, *-inā-*,<sup>191</sup> *-ā-* či *-ē-*, jenž se v přítomném kmenu vypouští (*-inā-* je redukováno na *-in-*), naopak v préteritálním kmenu se rozšiřuje o vkladný sufix *-j-*. V závislosti na tematickém sufixu se užívá se buď úplné

**klaus|īties** *v*<sub>3</sub> (~os, ~ies, ~ās; ~ījos) **1.** (*ko*) *poslouchāt (co/koho)* [...],

nebo částečně redukované (nulová koncovka pro 3. osobu obou čísel) sady osobních koncovek.

**grib|ēt** *v*<sub>3</sub> (~u, ~i, ~; ~ēju) **1.** *chtīt* [...]

<sup>190</sup> Ojedinelé též *-ū-* u slovesa *dabūt* „dostat, získat“.

<sup>191</sup> Spolu s výjimkou v podobě slovesa *zināt* „vědět, znát“.



U některých sloves se v přítomných tvarech může projevit hlásková změna, možné jsou i dublety.

**gu|lēt**  $v_3$  (~|u, ~li, ~; ~lējū) **1. spát** [...]

**sē|dēt**  $v_3$  (~žu<sup>192</sup>/~du [æ:], ~di, ~ž/~d [æ:]; ~dējū) **1. sedět** [...]

3.2.3.2.4.4 U trojice sloves atematických (*būt, iet, dot*) včetně jejich odvozenin používáme slovnědruhovou značku  $v_0$ , jejich supletivní tvary či kmeny zpracováváme jako odkazová hesla.

**būt**  $v_0$  (esmu [æ], esi, ir, esam [æ], esat [æ]; bijū) **1. být** [...]

**aizd|ot**  $v_0$  (~odu, ~od, ~od; ~evu) **1. (kam ko) půjčovat × půjčit (komu co)** [...]

**esmu** [æ] 1sg prez ↑ **būt**

**ej, ej–** prez. ↓ **iet**

**dev–** prét. ↓ **dot, doties**

3.2.3.2.4.5 Některá slovesa se časují podle paralelních paradigm, v běžně mluvené praxi se nezřídka objevují i paradigmata smíšená, hybridní, byť to není jazykovými příručkami doporučováno (LVG: 549, srov. čistě deskriptivistický přístup v Paegle 2006). Vzájemně si konkurující dublety, mají-li dostatečnou oporu v dnešním úzu,<sup>193</sup> budou v LČ, stejně jako je tomu ve většině překladových slovníků, postiženy; v exemplifikaci už je však zpravidla dána přednost četnějšímu tvaru. Paralelní formy budou řazeny podle frekvence, v případě nedostatku korpusových dat na základě informace ve výkladových slovnících, případně v jiných zdrojích (Paegle 2006).

**skrie|t**  $v_1$  (~nu / skreju, ~n / skrej, ~n / skrej; skrējū) **1. běžet** [...]

**pēt|īt**  $v_2/v_3$  (~tju/~u [æ:], ~t/~i, ~t/~a [æ:]; ~tju) **1. (ko) zkoumat, studovat (co), bádāt (nad čím)** [...]

3.2.3.2.5 Participia jsou slovesné formy mající zároveň povahu slovesnou a adjektivní (ohebná) či adverbialní (neohebná), tato dvojí povaha vede ke kom-

<sup>192</sup> Hláskově alternovaný kořen najde uživatel pod odkazovým heslem **sēž, sēž–** prez ↑ **sēdēt**

<sup>193</sup> Takto ignorujeme staré varianty sloves původně 2., dnes 3. konjugace, uváděné v LVG (550); nezaznamenávají je žádné synchronní slovníky ani korpusy. Migrace sloves z 2. do 3. konjugací třídy, doložitelná v lotyštině již od 18. století, je dosud neukončeným procesem (LVG: 550; Paegle 2006).

plikacím při gramatickém popisu. LČ chápe participia jako zvláštní třídu tvarů, což je zohledněno specifickou slovnědruhovou značkou *partic*.

V lotyštině rozlišujeme tyto druhy participií: vedle čtveřice ohebných (příčestí přítomného činného a trpného a minulého činného a trpného), pojednání již zčásti v kapitole o substantivech (3.2.3.1.4.8), jedno částečně ohebné (poloparticipium) a dvě neohebná (absolutiva).<sup>194</sup> Význam a funkce jednotlivých typů participií včetně různých možností překladu do češtiny<sup>195</sup> budou uživateli vyloženy v gramatickém přehledu.

3.2.3.2.5.1 Poloparticipium má pouze nominativní tvary, utvořené od infinitivního kmene rozšířeného o sufix *–dam–* příslušnou rodovo-číselnou koncovkou, kterou se shoduje se jménem. Na rozdíl od ohebných příčestí nemá poloparticipium tendenci lexikalizovat se, s jeho exemplifikací v rámci základového slovesa se tudíž počítá jen v případě ustálených obrátů:

**darjīt** *v*<sub>3</sub> (~u, ~i, ~a; ~īju) **1.** (*ko*) *dēlat* (*co*) [...] • ~i *ko* ~īdams, apdomā galu *ať dēláš cokoliv, myslí na konec* [...]

Uvádíme je naopak mezi morfologickými tvary u sloves s minulým kmenem zakončeným na *–t* či *–d*, dochází tu totiž k hláskové změně konce infinitivního kmene (*s* > *z*).

**atr|ast** *v*<sub>1</sub> (~odu, ~odi, ~od; ~adu; fut ~adīšu; partic ~azdams) **1.** (*ko* [*kur*]) *najīt, nalēzt* (*koho/co* [*kde*]) [...]

**mesties** *v*<sub>1</sub> (metos [æ], meties, metas [æ]; metos; partic mezdams [æ]) **1.** *vrhnout se* [...]

3.2.3.2.8.2 Absolutiva jsou neohebné participiální tvary adverbialní povahy: vyjadřují nejružnější okolnosti děje. Tvoří se od přítomného kmene formantem *–ot* či *–am/–ām*, u zvrtných sloves doplněných ještě o reflexivní koncovku

<sup>194</sup> Oba termíny uvedené v závorkách přejímáme od A. Erharta (1984: 133, 117, 119).

<sup>195</sup> Například možnost překládat lotyšské poloparticipium českým přechodníkem přítomným (participium minulé činné pak přechodníkem minulým) je vzhledem k příznakovosti přechodníků v současné češtině omezena pouze na specifický typ kontextu (např. odborné či umělecké texty) a měla by být uživateli nabídnuta pouze jako alternativa k bezpříznakovému ekvivalentu. Uvádět k lotyšským participiím jako ekvivalenty české přechodníky přímo v hesláři se nám tak jeví jako chybné a zavádějící lexikografické řešení.

(*–oties*; *–amies/–āmies*). Běžnější je prvně jmenovaný tvar (LVG: 591–592),<sup>196</sup> mající nejuniverzálnější možnosti užití: téměř ve všech případech může konkurovat druhému neohebnému tvaru i poloparticipiu; čistě synonymické prostředky to však nejsou.<sup>197</sup>

Některé tvary neohebných participií se adverbizovaly či prepozicionalizovaly, takové zahrnujeme do hesláře jako samostatná lemmata s příslušnou slovnědruhovou zkratkou. Na rozdíl od jiných překladových slovníků však osamostatňujeme i další participia, jež se stala součástí ustálených obrátů.<sup>198</sup>

**sakot** partic: • ar vienu vārdū s. *jedním slovem* • citiem vārdiem s. *jinými slovy [řečeno]* • patiesību (taisnību) s. *popravdě řečeno* • godīgi (atklāti) s. *upřímně řečeno* • pareizāk s. *lépe (přesněji) řečeno* • īsāk s. *stručně řečeno, zkratka, ve zkratce* • tā s. *řekněme, jak se říká, takřkajíc* || též ↑sacīt<sup>199</sup>

**sākot** partic: • s. ar (*ko*), s. no (*kā*) (ve funkci nevlastní předložky) *od (koho/čeho), počínaje (kým/čím)* • plānots, ka, s. no 2007 gada, fondā būs pieejami 8 miljoni latu *plánuje se, že od roku 2007 bude ve fondu k dispozici 8 milionů latů* || též ↓sākt

Úzus: V LVK se tvar sākot objevuje 466×, z toho ve spojení s předložkami ar 217× a no 149×. Téměř v 80 % případů tak plní funkci sekundární předložky (též → **Gramatika: s. 1060**).

<sup>196</sup> Frekvenci užití jednotlivých typů přičestí nelze bohužel ověřit na korpusových datech. LVK uvádí tyto údaje:

- celkový počet participií [tag="v..p.\*"]: 244 013;
- z toho ohebných [tag="v..pd.\*"]: 207 214;
- částečně ohebných [tag="v..pp.\*"]: 2687;
- neohebných [tag="v..pu.\*"] 34 112 – ty můžeme pomocí pozitivního filtru *.\*ot(ies)?*, resp. *.\*m(ies)?* dále vydělit na obě podskupiny, jež jsou ve výrazném nepoměru: 33 178 (= 97,3 %) × 929. Skupinu neohebných participií končících na *–a/ām(ies)* jsme podrobili manuální analýze, přičemž skutečných neohebných participiálních tvarů se v daném seznamu konkordancí vyskytuje pouhých 147 (sic!), tj. 15,8 % – docházíme tak k závratné, více než 84% chybovosti (srov. oficiálně uváděnou *úspěšnost* morfologického taggeru lotyštiny SemTi-Kamols 93,6 % pro celý morfologický tagset, respektive 97,9 % pro určování slovních druhů: Paikens – Rituma – Pretkalniņa 2013). Drtivou většinu příkladů (711, tedy 76,3 %) tu tvoří homonymní tvary 1. osoby plurálu indikativu přítomnosti, dále jsou tu špatně rozpoznané a otagované formy dativu singuláru (13) a plurálu (15), částečně ohebného participia (24), 1 výskyt přičestí minulého činného, 12 zkratk, 2 výskyty zjevně chybné a 4 sporné případy, jež však hledaným participiovým tvarem být nemohou.

Dílčí frekvenční statistiku participií, založenou na analýze jedné obsáhlejší povídky Regīny Ezerové, podává LVG (568): přičestí minulé činné 348, přičestí minulé trpné 147, přičestí přítomné trpné 99, poloparticipium 99, absolutiv *–ot(ies)* 44, přičestí přítomné činné 29, absolutiv *–a/ām(ies)* 3.

<sup>197</sup> Přičemž sama LVG: 593 doznává mlhavosti kritérií sloužících k sémantické distinkci užití toho či onoho prostředku, záleží především na mluvčím a jeho preferencích.

<sup>198</sup> Níže uvedené participium *sakot*, doložené v LVK 419×, je v LVG (587) uváděno v obrazech *tā s.* (v LVK frekvence 20), *pareizāk s.* (41), *ar vienu vārdū s.* (68), *patiesību s.* (26), vedle nich nacházíme v LVK ještě další kandidáty na zařazení do LČ: *godīgi s.* (39), *taisnību s.* (20), *īsāk s.* (16), *atklāti s.* (10), *precīzāk s.* (6). Námi navrhovaná struktura hesla odráží frekvenci jednotlivých spojení.

<sup>199</sup> Dvojtečka hned za slovnědruhovou zkratkou naznačuje restriktce typického použití lexému, odkaz na konci pak upozorňuje na možnost užití tvaru i v jeho primární, tj. nedesémantizované podobě.

V žádném z nám známých překladových slovníků není ani jedno z těchto dvou participií uvedeno jako samostatné heslo, běžná praxe je zařazovat tyto tvary pod lemma základového slovesa, které však nemusí být vždy zjevné (hlásková alternace *c–k* u prvního participia). Jejich explicitním zařazením do hesláře, mimoto takto bezprostředně za sebou, je eliminováno riziko uživatelské mylné interpretace.

Nedesémantizované tvary absolutiva se mohou objevit v exemplifikaci základového slovesa, případně sloves jiných (*verba sentiendi, declarandi* aj.):

**justies** *v*<sub>1</sub> (jūtos, jūties, jūtas; jutos; fut jūtišos) **1.** (*kā, kāds*) *cítīt se* (*jak, jaký*) [...] • *viņš jūtas parādā esam cītī se būt dlužný, má pocīt dluhu*

**sajcīties** *v*<sub>3</sub> (~kos, ~kies, ~kās; ~cījos) **1.** (*ve spojení s neohebným přičestím*) *tvrdīt* (*co / že*) • *viņš ~kās esam slims tvrdī, že je nemocný* • *viņas ~kās neko nezinām par šo lietu tvrdī, že o této věci nic nevēdī* [...]

### 3.2.3.3 Gramatická charakteristika konjunkcí v LČ

Lotyšské spojky pro účely zpracování v LČ dělíme, v souladu s aktuální lotyšskou gramatikou (LVG: 641–643), do tří skupin: na jednoslovné a více-slovné, z nichž vydělujeme ještě specifickou skupinu spojek párových, jež se ve větě vyskytují odloučeně, přesto však tvoří nedílnou lexikální jednotku. Běžně se v lotyšských slovnících jakožto samostatná hesla zpracovávají pouze jednoslovné spojky; analytické tvary bývají většinou zahrnovány pod lemma jednoho či obou komponentů (patřících však nežádka k jinému slovnímu druhu) jako kolokace, neřkuli frazémy, samostatně se objevují řídce, nahodile.<sup>200</sup> V LČ se pro ně počítá s vlastními hesly, nehledě na jejich složitost; primárním kritériem je míra ustálenosti a soběstačnosti daného lexému. Takto např. nepovažujeme níže

<sup>200</sup> Srov. překvapivě (?) shodné zařazení lemmat *tā ka* i *tā kā* do heslářů LR, LA i LP, nikoliv už však jiných, neméně potřebných. Za zmínku stojí též řazení obou hesel podle tzv. *letter-by-letter* principu (Svensén 2009: 369–370):

LR: ..., *taisnzarņi, tā ka, tā kā, tak, ...*;

LA: ..., *taka, tā ka, tā kā, takelāža, ...*; stejně tak LP (1341), jen navíc s homonymním heslem *tā kā* (které je do samostatného hesla vydělováno zbytečně, jde o zjevnou polysémii).

V LČ se počítá s řazením podle tzv. *word-by-word* principu, tj. víceslovná lemmata budou bezprostředně následovat po jednoslovných: ..., *ta, TA* [zkr.], *tā, tā arī, tā ka, tā kā, tabaka, ...* Výhodou našeho řešení je možnost okamžité detekce všech víceslovných jednotek, jejichž je jejich první komponent součástí.

uvedenou víceslovnou spojkou *kaut gan* (spolu se synonymickými spojkami *kaut arī* či *lai gan* a *lai arī*)<sup>201</sup> za realizaci jednoslovné konjunkce *kaut* (ani partikule *gan* či *arī*),<sup>202</sup> ale za svébytný analytický tvar, vyžadující si specifický popis, tudíž i vlastní heslo; podobně nelze směšovat ani zesilovací částici *gan* s párovou spojkou *gan–gan*.

**kaut gan** konj *i když, přestože, ačkoli[v], třebaže* • k. g. nebija vēls, ārā jau kļuva tumšs *i když nebylo pozdě, venku už se setmělo* • k. g. ļoti gribējās, es tomēr nekā neteicu *ačkoliv se mi moc chtělo něco říct, přesto jsem nic neřekl* • bieži esmu neizlēmīgs, k. g. no malas tā nešķiet *často bývám nerozhodný, i když zvnějšku se to nezdá*

**gan ..**<sup>203</sup> **gan** konj *jak... tak [i], i... i* • g. mēs, g. viņš *jak my, tak i on* • g. šā, g. tā *tak i tak* • man patīk dzert g. tēju, g. kafiju *rād piju jak čaj, tak kávu*

**ne tikai .. bet arī** konj *nejen[om]... ale i* • saldumus mīl ne t. bērni, b. a. pieaugušie *sladkosti milují nejen děti, ale i dospělí*

V případě polysémie přihlížíme i k sémantické stránce spojek: uvádíme typ vztahu mezi větami či větnými členy, případně doplněný ještě o metajazykovou glosu.

**bet**<sup>1</sup> konj **1.** (odporovací) *ale, zato, a* • mās, b. spēcīgs *malý, ale (zato) silný* • es gribu gulēt, b. nevaru aizmigt *chci spát, ale (a) nemůžu usnout* • mēs iesim, b. tu paliksi šeit *my půjdeme, ale (a) ty zůstaneš tady* • Latvijas valdībā veselības ministra pienākumi uzticēti politiķim, kas pēc savas izglītības nav ārsts, b. gan ekonomists. (důraz na protiklad) *V lotyšské vládě jsou (byly) povinnosti ministra zdravotnictví svěřeny politikovi, který je vzděláním nikoliv lékař, ale ekonom.* • Dažreiz par to iedomājos. Bet tad tas sapnis laikam vairs nerādītos. *Občas jsem si to představoval. Ale pak by se ten sen už nesplnil.* || **2.** (slučovací, liší-li se podměty dvou vět) *a* • mans tēvs strādā par ārstu, bet māte par skolotāju *můj otec pracuje jako lékař, matka jako učitelka* || jako součást víceslovných spojek ↓ ne .. bet, ne tikai .. bet arī, ne vien .. bet arī, nevis .. bet

**bet**<sup>2</sup> m<sub>0</sub> <hov> (háček, zádrhel) *ale* • tas būtu labi, taču ir viens b. *to by bylo dobré, má to však jedno ale* • nekādu b. *žādnē ale!*

Roli spojovacích výrazů plní rovněž některá vztažná příslovce (*cik, kā, kad, kādēļ* a *kur*) a vztažná zájmena (*kas, kurš* a *kāds*, respektive všechny jejich pádové formy), mající na rozdíl od konjunkcí rovněž větněčlenskou platnost.

<sup>201</sup> jež by měly být pořešeny takto, s ohledem na korpusová data (odkazová/samostatná hesla?)

<sup>202</sup> Odkaz (též ↓ *kaut gan*, resp. též ↓ *kaut arī*) uvedený na konci hesla *gan* (resp. *arī*) je čistě pragmatickým prostředkem sloužícím k orientaci uživatele, jenž správně nerozpoznal víceslovný lexém a snaží se jej dešifrovat podle druhého komponentu.

<sup>203</sup> Párové spojky jsou od sebe v hesláři LČ odděleny dvojtečkou, používanou v lotyštině pro označení elipsy, český ekvivalent, je-li rovněž vícedílný, pak trojtečkou. (Viz heslo *divpunkte* v kap. 3.2.7.)

Většina těchto lexémů není v LČ chápána jako homonymní, ale polysémní: např. vztažná zájmena jsou spolu s tázacími zpracována v rámci jednoho lemmatu, nicméně pod různými, oddělenými významy. Za homonyma považujeme pouze ty případy lexémů zjevně se od sebe lišících slovním druhem.

**kad**<sup>1</sup> adv 1. (tázací) *kdy* • k. tu brauksi / būsi mājās? *kdy přijedeš / budeš doma?* • k. [tad] tas bija? *kdy (kdypak) to bylo?* • k. beidzot klūs siltāks? *kdy se konečně oteplí?* || 2. (vztažné) *kdy* • nezinu, kad nākšu *nevím, kdy přijdu* • diez vai k. redzēsīties *sotva (těžko) se někdy (kdy) uvidíme* •

**kad**<sup>2</sup> konj 1. (uvozuje vedlejší větu časovou) *když* • pēc tam, k. *potom co, poté co* • jau krēsloja, k. viņš devās mājās *už se stmívalo, když se vydal domů* • kad aizsūtīju vēstuli, atbildi nesaņēmu *když jsem poslal dopis, odpověď jsem nedostal* || 2. (... přívlastkovou n. doplňkovou) *když* • atnākusi vasara, kad naktis pavisam īsas • vajadzētu pajautāt, kad pienāk autobuss || 3. (...podmínkovou) *když* • kādas mežā sēnes, kad lietus nav lijis *jak můžou být v lese houby, když nepršelo* || 4. <hov> ↑ka<sup>1</sup> || 5. <zast> ↓kaut<sup>1</sup> || ♦ k. jau, tad jau *když už, tak už*

**kad**<sup>3</sup> part <hov.> *kdyby [tak], kéž by* • k. tikai neuznāk lietus! *kdyby jen nepršelo!* • k. tik vien mums kāds palūgtu *kdyby nám tak někdo pomohl*

### 3.2.4 Informace sémantická

3.2.4.1 Nositeli sémantické informace v naší databázi jsou dva vzájemně provázané elementy: Význam (viz 2.4.5) a jeho bezprostřední potomek Ekvivalent (2.4.6).<sup>204</sup> Hierarchický poměr příhodně odráží dvě odlišné, nicméně na sebe navazující fáze při budování heslové stati polysémního hesla: prvotní rozvrstvení sémantického struktury lotyšského lexému na jednotlivé významy<sup>205</sup> a následné hledání českých protějšků (srov. lišení překladového ekvivalentu a překladového významu jako dvou aspektů ekvivalence u A. Jarošové [2008a]). U hesel monosémních je situace o poznání jednodušší, neboť tu odpadá první jmenovaná fáze. Přesto i zde zůstává nutnost spojit český protějšek s lotyšským lemmatem prostřednictvím elementu Význam (viz též obr. 5), byť tento zůstane neočíslován – v dvoujazyčném slovníku je totiž význam daného lexému ve vý-

<sup>204</sup> Ten může fungovat rovněž jako potomek atributů Kolokace a Frazém, v obou případech zůstává nositelem sémantické informace.

<sup>205</sup> Či spíše jejich potenciály (Hanks 2000): „There is no direct route from the corpus to the meaning. [...] The corpus contains traces of meaning events; the dictionary contains lists of meaning potentials. Mapping the one onto the other is a complex task, for which adequate tools and procedures remain to be devised.“

Srov. též Zgustovo pojetí vrcholů v rámci kontinua realizací: „[...] we might consider single senses are characteristic peaks in the continuum of possible applications.“ (Zgusta 1971: 70) „[...] the applications form a continuum with observable peaks but unclear boundaries.“ (ibid.: 74)

chozím jazyce popisován prostředky jazyka cílového.<sup>206</sup> Platí rovněž, že významové členění hesla je primárně určováno jazykem cílovým (Čermák 1995d: 232, 244), překladový slovník by neměl být „překladem“ slovníku výkladového (Svensén 2009: 278; srov. též např. Adamska-Sałaciak 2006: 96–97, Sekaninová 1993: 46–49, Jarošová 2000 aj.)

Toto specifikum kontrastivní lexikografie staví do centra její pozornosti pojem ekvivalence, kterou F. Čermák (1995d: 238) charakterizuje jako „vztah mezi lexémem v určitém významu výchozího jazyka a formou cílového jazyka založený na jejich (téměř) identickém, resp. vysoce podobném významu a funkci“. Při hledání takového vztahu naráží lexikograf na dvojí překážku: jednak se musí vyrovnat s obecnou vlastností jakékoliv kombinace jazyků, a sice s jejich anizomorfismem (např. Zgusta 1971: 294n.); jednak s dvojí povahou lexému: abstraktní, systémovou oproti konkrétní, textové (Jarošová 2000: 19, 2008a: 105).

Nejčastěji bývá rozlišován trojí druh ekvivalence,<sup>207</sup> z nichž oba krajní póly – ekvivalence úplná a nulová – jsou poměrně řídké. První případ, kdy si výraz ve zdrojovém a v cílovém jazyku plně odpovídají, je omezen de facto pouze na odbornou terminologii, přesně definovanou (často i standardizovanou) a účelově monosémní. Druhý případ představuje situaci, v níž výrazu ve zdrojovém jazyku neodpovídá žádný sémanticko-funkční protějšek v jazyce cílovém, se rovněž týká specifické části lexikonu: jde většinou o nejrůznější kulturně podmíněné (*culture-specific*<sup>208</sup>) reálie téměř výhradně substantivního charakteru (např. pokrmy, nejrůznější nástroje, společensko-politické a duchovní jevy apod.); tyto lakuny je však pokaždé možné (byť ne vždy snadno) zacelit alternativním způsobem: opisným ekvivalentem, popisujícím daný lexém z hlediska vnějšího

---

<sup>206</sup> Srov. Svensén 2009: 253 či Kopeckij 1974: 35: „Není náhodné, že jedním z nejefektivnějších způsobů vymezení významu a sémantiky slova s jeho sémantickými množiteli se stává jeho překlad.“ Podobně se k významu jakožto prosté funkci spojující slovo v jazyce A s jiným slovem v jazyce B (jenž funguje jako metajazyk) staví J. Clear (1996: 270).

<sup>207</sup> Srov. však též jiné pojetí, např. Sekaninová 1993: 58–79.

<sup>208</sup> Srov. Zgustova „culture-bound words, i.e. lexical units whose denotata or designate simply do not exist in another culture“ (1971: 114, dále též 294, 324), od nichž však odlišuje „onomasiological gaps“ (ibid.: 325).

a/nebo funkčního, přejímkou, redukcí některých významových rysů, případně ještě jinými řešeními.<sup>209</sup>

Mezi těmito dvěma krajními body se rozprostírá široká, diverzifikovaná škála případů ekvivalence částečné, s rozmanitou mírou totožnosti srovnávaných výrazů.<sup>210</sup> S tou se lexikograf při tvorbě slovníku potýká nejčastěji. Nesymetrický vztah mezi lexémy obou jazyků může kompenzovat různými prostředky.<sup>211</sup> Nejběžnějším z nich je glosa, sloužící ke zpřesnění kontextu, v němž se daný ekvivalent může vyskytnout (a tedy i naznačující jeho elementární kolokabilitu), v případě vícerych synonymických protějšků k jejich vzájemné distinkci. Glosy by měly být co nejlapidárnější, je třeba dbát rovněž na jejich konzistenci u analogických případů; využít lze prostředků jak paradigmatických (synonymie, též slovníky často nevyužívaná opozitnost, dále hyperonymie/hyponymie, meronymie), tak syntagmatických (upřesnění aktantů, okolností, vztahů apod.).

V LČ jsou glosy, psané výhradně v češtině, situovány většinou před ekvivalent, jakožto prostředek primárně desambiguační: vycházíme z předpokladu, že uživateli mohou sloužit jako jakési návěstí, pomáhající mu snáze se orientovat ve struktuře heslové stati. Glosa ekvivalentu předsazená mu jasně naznačí kontext, a pokud ten nekoresponduje s textovou realizací lexému, může uživatel přejít k dalšímu ekvivalentu, případně k jinému významu uvozenému arabskou číslicí (ta sama o sobě žádnou distinkční charakter nemá, pouze vyděluje konkrétní ekvivalent[y] od ostatních, nesynonymických). Proti doporučení

---

<sup>209</sup> Přibližnost či částečnost ekvivalentu lze explicitně naznačit též vyhovujícím grafickým prostředkem (např. ČLN v tomto ohledu využívá znak  $\cong$ , srov. též Nikulceva 2013: 216).

<sup>210</sup> Ve zvláštním případě tzv. polyekvivalentace vidí F. Čermák pouhou amplifikaci případu částečné ekvivalence, „komplexní souvztažnost[i] mezi dvěma skupinami lexémů“ (Čermák 1995d: 238). Viz též Jarošová (2000: 20), citující charakteristiku takového vztahu z pera F. J. Hausmanna („ein gewaltiges many-to-many Geflecht, das nicht mehr überschaubar ist“) a tlumočící jeho skeptický postoj: „If we begin to test the mutual insertability of both paradigm units in real contexts we are then shifting to the level of contrastive textology, where the systematization and formulation of equivalent relations is, despite the help of computers, according to Hausmann, impossible.“ (ibid.)

<sup>211</sup> Jarošová 2008b: 74: „[...] the process of compiling of a bilingual dictionary entry can be described as building up artificial semantic equalities from semantically anisomorphic entities (words and fixed expressions) by using specific lexicographical tools: [...] These compensatory tools signalize the conditions (verbal or situational context) under which semantically non-identical entities (if considered in isolation) can be equal.“ Srov. též Svensén 2009: 261.



A. Adamské–Šalacziakové (2006: 89–90) psát glosy za ekvivalent, protože tak mimo jiné vytvářejí smysluplné kolokace (např. *deska* (stolu), *česat* (vlnu, len)), lze namítnout, že transformace z opačné struktury (stolu) *deska* či (vlnu, len) *česat* je operace natolik elementární, že by uživateli neměla činit žádné obtíže.

**āzis** m<sub>2</sub> (āža) **1.** (samec kozy, popř. též jiných býložravců) *kozel* • stirnu ā. *srnec* || **2.** (tělocvičné nářadí) *koza* || **3.** (podstavec k řezání dřeva) *koza* • zāģēt malku uz āžiem *řezat dřevo na koze* || ♦ āža bārda *kozí bradka* ♦ grēka ā. *obětní beránek* ♦ dīrāt āži <hov> *blejt, zvracet* [...]

**blīkšķļēt** v<sub>3</sub> (~; ~ēja) (hrom, výstřel, dveře ap.) *prásknout, třísknout, bouchnout*

**pusvienpadsmit** num (jen v časových údajích: hodiny) *pūl jedenácté* • pulkstenis ir p. *je pūl je-denácté* • atgriezties mājās ~os vakarā *vrátit se domů v pūl jedenácté večer*

Stranu napravo od ekvivalentu ponecháváme pro doplňkové informace, zpravidla – leč ne výhradně – encyklopedického rázu.

**bermontiāde** [o] f<sub>5</sub> <hist> *bermontiāda* (vojenské akce r. 1919 vedoucí k vytlačení armády plukovníka Bermondta-Avalova lot. a estonskými vojsky z území Lotyšska; též Bermondtova aféra)

**Neptūns** m<sub>1</sub> **1.** <astr> *Neptun* • ~a pavadoņi *měsíce Neptunu* || **2.** <mytol> *Neptun* (římský bůh moře)

**klon|s**<sup>2</sup> [uo] m<sub>1</sub> **1.** *podlaha* (obv. z udusané hlíny či vydlážděná) **2.** *mlat, humno* [...]

**ābe|ce** f<sub>5</sub> (Gpl ~ču) **1.** *abeceda* [...] || ♦ ~ces patiesība *základní pravda* (všeobecně známá, ne-zpochybňovaná) [...]

**sievasbrālis** m<sub>2</sub> (~ja) *švagr* (z manželčiny strany)<sup>212</sup>

Obdobně je zapotřebí vysvětlující glosou opatřit také synsémantika; primárním úkolem je v takovém případě stručně popsat jejich funkci a tento popis doplnit o vhodný příklad. Typ funkční glosy je v TLexu pro větší přehlednost odlišen od atributů Kontext (eventuálně i Kontext\_před/Kontext\_za – viz 2.4.5.6 a 2.4.6.4) jako samostatný atribut (Funkce – viz 2.4.5.7), formálně však zůstává stejný.

**bet**<sup>1</sup> konj **1.** (odporovací) *ale, zato, a* [...] • Latvijas valdībā veselības ministra pienākumi uzticēti politiķim, kas pēc savas izglītības nav ārsts, b. gan ekonomists. (důraz na protiklad) *V lotyšské vládě jsou (byly) povinnosti ministra zdravotnictví svěřeny politikovi, který je vzděláním nikoliv lékař, ale ekonom.* • Dažreiz par to iedomājos. Bet tad tas sapnis laikiem vairs nerādītos. *Občas jsem si to představoval. Ale pak by se ten sen už nesplnil.* || **2.** (slučovací, liší-li se podměty

<sup>212</sup> Jiné je *vīrabrālis* „švagr (z manželovy strany)“. K příkladům konvergence a divergence mezi výchozím a zdrojovým jazykem viz např. Svensén 2009: 259–260.

dvou vět) *a • mans tēvs strādā par ārstu, bet māte par skolotāju mūj otec pracuje jako lékař, matka jako učitelka*

**visu labu** interj (pozdrav při loučení, spolu s přáním všeho dobrého) *nashledanou, naviděnou || hodně štěstí (zdaru), měj (mějte) se dobře (hezky)*

**jā–** prefix debitivu, ve spojení s 3. osobou přítomného času vyjadřuje nutnost, nezbytnost, potřebu apod., např. *man [ir] jāiet* („musím jít“), *tev ir mani jāapciemo* („musíš mě navštívit“), *viņai būs jāatrod citu vietu* („bude muset najít jiné místo“) → **Gramatika: s. 1052**

Na upřesnění ekvivalentu, respektive distinkce konkrétního významu (či funkce) a úzu daného lexému, se však mohou podílet také další typy informací, o nichž tu už byla a ještě bude řeč, počínaje charakteristikou fonologickou a konče exemplifikací, ilustrující konkrétní realizaci lexému v textu. Často bývá podceňována a zanedbávána pragmatická složka lexému, již je vhodné rovněž zachytit: ať už odpovídajícím klasifikátorem, stručnou glosou anebo explicitním popisem konotativních, evaluativních, emotivních aj. specifík výrazu (viz dále 3.2.7). Se všemi těmito externími vazbami musí být sémantická část, tvořící středobod heslové stati, v souladu (Čermák 1995d: 242).

**gnīd|a** 1. f<sub>4</sub> (vajíčko vši) *hnida* 2. m/f<sub>4</sub> (Dsg m ~am, f ~ai) <hanl> (člověk) *všivák, prašivec, hnida, veš • g. tu tāds! ty seš ale všivák! všiváku jeden!*

**pār|is**<sup>1</sup> m<sub>1</sub> 1. (lidí, zvířat) *pār, dvojice [...]* 2. (předmětů) *pār • apavu / zeķu p. pār bot / ponožek [...]*

**pāris**<sup>2</sup> num<sub>0</sub> (dva; několik) *pār • pateikt p. vārdi říci pār slov • pēc p. dienām za pár dnů • p. reižu párkārt, nēkolikrāt [...]*

3.2.4.2 Kromě vlastního protějšku uvádí LČ v některých případech i sekundární ekvivalent, realizovaný atributem Latina (viz 2.4.6.2). Nejčastěji jde o latinské názvy taxonomických jednotek (srov. např. Zgusta 1971: 331–332, Čermák 1995d: 245), případně jiné cizojazyčné názvy zejména proprií. Lze je vnímat jako třetí členy triád mezijazykových synonym (Sikogukira 1994, Murphy 2003: 146, 155),<sup>213</sup> jež mají daný ekvivalent snadněji a přesněji ztotožnit s výchozím výrazem, případně též posloužit k desambiguaci (*rozpuk* ‚rozpučení × rostlina‘); fungují většinou jako extenze diatechnických klasifikátorů.

<sup>213</sup> Srov. koncept mezijazykové homonymie (případně paronymie) v podobě tzv. falešných přátel, *faux amis*.

**velnarutks** [æ] m<sub>1</sub> <bot> *rozpuk* (Cicuta)

**Ahillejs** m<sub>1</sub> <mytol> *Achilleus* • ~a cīpsla <anat> *Achillova šlacha* (tendo Achilles), *achilovka* [...]

**Veimārja** f<sub>4</sub> *Výmar* (Weimar) • ~as republika <hist> *Výmarská republika*

Podobně identifikační funkci mají sekundární protějšky také v případě lexémů nahrazovanými běžně zkratkami a grafickými symboly.

**kadmijls** [æ] m<sub>1</sub> <chem> *kadmium* (Cd) • niķeļa--a akumulators <elektr> *nikl-kadmiový akumulátor* (NiCd)

**diēzs** m<sub>1</sub> <hud> *křížek* (#)

**promīle** (Gpl ~ju) f<sub>5</sub> *promile* (‰) • 1,5 ~les alkohola asinīs *jeden a půl promile alkoholu v krvi*

Pokud to lze, opatřujeme lotyšská křestní jména českými protějšky, přičemž tertium comparationis je společný základ obou proprií. V určitých kontextech (níže uvedený příklad biblismu; překlad dětské knihy, kdy se překladatel rozhodne jména hlavních hrdinů počestit, aj.) totiž může být vhodné nezachovat originální podobu jména.

**Uģlis** m<sub>2</sub> (Vsg ~il) lot. mužské křestní jméno (≡ č. Hugo)

**leva** f<sub>4</sub> (Vsg ~!) lot. ženské křestní jméno (≡ č. Eva) • Ādams un Ie. <bibl> *Adam a Eva*

3.2.4.3 Samotný proces ekvivalentace, tedy hledání vhodného protějšku v cílovém jazyce k lexému jazyka výchozího, vychází dnes oproti dřívější situaci daleko více z korpusových dat; především paralelní korpusy nabízejí lexikografovi cenné kandidáty na ekvivalenty. Stejně tak lze čerpat z dvoujazyčných slovníků opačného směru (viz praktická ukázka v 2.3.2.2) a zajisté je vhodné přihlídnout i k rozdělení sémantické struktury ve výkladových slovnících.<sup>214</sup> Základní, dichotomické členění ekvivalentů, opírající se o opozici *langue–parole*, bylo popsáno L. Zgustou (1971: 319n.): rozlišuje mezi ekvivalenty překladovými (*translational, insertable*), které jsou v cílovém jazyce ustálenými lexikálními jednotkami a mohou být použity v konkrétním textu bez dalších

<sup>214</sup> Platí však, že významové členění hesla je primárně určováno jazykem cílovým, nikoliv výchozím (Čermák 1995d: 232, 244), překladový slovník by neměl být „překladem“ slovníku výkladového (Svensén 2009: 278); srov. též např. Adamska-Sałaciak 2006: 96–97, Sekaninová 1993: 46–49, Jarošová 2000 aj.

úprav, a ekvivalenty vysvětlujícími (*explanatory, descriptive*), jež jsou libovolnými prostředky cílového jazyka sloužícími k převodu určitého lexému v daném textu.<sup>215</sup>

S pokusem o rozšířenou koncepci, z níž vycházíme i my při přípravě LČ, přišla A. Adamská–Saľacziaková (2006: 103n.). Ekvivalenty jsou klasifikovány do čtyř tříd:

- ekvivalenty vysvětlující (*explanatory, též descriptive*) jsou parafrázemi lexémů zdrojového jazyka v jazyce cílovém, nejčastěji jejich definicemi. V cílovém jazyce nebývají ustálené. LČ se k nim uchyluje zejména v případě reálií, jež jsou v českém prostředí neznámé.

**ačkups** m<sub>1</sub> <folk> lot. lidový tanec

**LETA** [leta] f<sub>4</sub> lot. národní tisková agentura

**Salaspilis** f<sub>6</sub> město ve stř. Lotyšsku; správní středisko stejnojmenného kraje

- ekvivalenty kognitivní (*cognitive, též semantic, systemic, prototypical, conceptual, decontextualized, notional*) mají všeobecnou, languovou povahu, pokrývajíce především centrální, prototypické významy lemmat. Od předešlé skupiny se odlišují svou ustáleností v cílovém jazyce. Kognitivních ekvivalentů je v každém dvoujazyčném slovníku zdaleka nejvíce.

**piecdesmit** num<sub>0</sub> *padesát* [...]

**ābe|ce** f<sub>5</sub> (Gpl ~ču) **1.** *abeceda* [...] || **2.** *slabikář* [...] || **3.** (v nějakém oboru) *abeceda, základy* [...]

- ekvivalenty překladové (*translation[al], též insertable*) odpovídají Zgustově pojetí: jsou to autenticky doložené protějšky v konkrétních textech, tedy fungující v parolové rovině. Ty už z podstaty věci nepodchytí žádný, ani

<sup>215</sup> Viz též Zgustův okřídlený výrok: „[...] equivalents that are both translational and explanatory are the bilingual lexicographer’s best friends“ (Zgusta 1984: 148). Ideálně by slovníky měly poskytovat „real lexical units of TL which, when inserted into the context, produce a smooth translation“ (ibid.: 147), tedy dokonalé lexikální protějšky, což je také namnoze očekávání většiny uživatelů, samotným Zgustou hodnocené jako „naprosto přirozené“ (ibid.).

Srov. též modifikace Zgustovy základního rozdělení, lišící se mnohdy jen terminologicky, shrnuté Adamskou–Saľacziakovou (2006: 101).

sebeobsáhlejší slovník bezezbytku,<sup>216</sup> může pouze naznačit bohatost překladových možností vhodně zvoleným příkladem, ideálně převzatým (případně modifikovaným) z paralelního korpusu.

**diezin vai; diez vai** part 1. *sotva, těžko, stěží* [...] • Varbūt mums kaut ko vajag? – D. v. *Nepotřebujeme něco? – Nemyslím.* [...]

- ekvivalenty funkční (*functional*, též *situational*, *communicative*, *discourse*, *dynamic*) se s výrazy ve zdrojovém jazyce neshodují svou formou (viz též poslední příklad, kdy lotyšské partikuli odpovídá finitní slovesný tvar), plní však stejnou funkci. Vznikají často v důsledku syntaktické transformace, typické jsou ale rovněž pro idiomatická spojení.

**gribļēties** v<sub>3</sub> (~as; ~ējās) 1. (*kam infin*) *chtít se* (*komu infin*) • man ~as ēst / dzert *mám hlad / žízeň* • man ~as gulēt *chce se mi spát* [...] • dažkārt tik ļoti ne~as nekur iet *občas se člověku vůbec nechce nikam chodit (jít)* 2. (*kam ko*) *mít chuť* (*na co*) • man ~as kafiju *mám chuť na kávu, dal bych si kávu* [...] ♦ jo ēd, jo ~as *s jídlem roste chuť*

**beša:** ♦ palikt b. *zůstat s prázdnou, ostrouhat* ♦ atstāt b. (*ko*) *nechat na holičkách (koho)* [...]

3.2.4.4 Stejně jako je předpokladem pro zařazení lotyšského lexému do hesláře jeho opakovaný výskyt, je i evidence českého ekvivalentu podmíněna jeho rekurentností.<sup>217</sup> Od svojských překladatelských řešení je třeba odhlédnout, byť by byla sebelepší a sebeoriginálnější (Čermák 1995d: 239). To už přecházíme do fáze finálního výběru ekvivalentů a jejich následného usouvztažnění. Kromě frekvenčního hlediska se řídíme rovněž stupněm ekvivalence. Ekvivalenty nesynonymické konstituují samostatné významy (Ďurovič 1953), oddělené v LČ dvojitou svislicí; ty však nemusí být nutně číslovány (např. jde-li o pouhé

<sup>216</sup> „A bilingual dictionary could never give all type T equivalents of a given SL item, because it is impossible to predict all the contexts in which the item may occur.“ (Adamska–Saľacziak 2010: 398)

<sup>217</sup> „What matches a translation unit with its equivalent in a given target language is not some abstract property of the language system in general or the systems of two specific languages but the continuous practice of generations of translators.“ (Teubert 2002: 211–212)

významové nuance,<sup>218</sup> je-li zapojen historický aspekt, k odlišení odborného a laického úzu, v některých případech konverzivnosti apod.).

**abonēt** [o] v<sub>2</sub> (~ju, ~, ~; ~ju) (*ko*) (tiskoviny, divadlo, služby ap.) *předplácet* [si] × *předplatit* [si], *abonovat* (co) || (tiskoviny též) *odebírat* (co) || (služby též) *objednávat* [si] × *objednat* [si] (co) || (k odběru) *přihlásit se* (k čemu/kam) • a. avízi / žurnálu *předplatit si noviny / časopis*

**asi|ns** f<sub>6</sub> (~ns, Gpl ~nu) *krev* • *izliet* ~nis (*par ko*) *prolít krev* (za koho/co) • *nodot* ~nis *darovat krev* • ~ns *lāse kapka krve* [...] || <anat, med> (v odborném diskurzu převažuje singulár) • ~ns *analīze rozbor krve* • ~ns *saindēšanās otrava krve* • ~ns *pārliešana transfúze krve* • ~s *grupa krevní skupina* || ♦ *sarkans kā ~nis červý (rudý) jako krev* [...]

**Kauņa** f<sub>4</sub> *Kaunas* (město v Litvě) || <hist> *Kovno*

**aitasgalvja** m/f<sub>4</sub> (Dsg m ~am, f ~ai) <hanl> *hlava skopová* || (muž) *hlupák, beran* || (žena) *hlupač-ka*<sup>219</sup>

Rovněž tak je obhajitelné jisté zjednodušení a redukce sémantické struktury u některých specifických tříd slov.

**Milda** f<sub>4</sub> <hov> Pomník svobody v Rize; socha ženy na tomto pomníku

**Ogre** f<sub>5</sub> město ve stř. Lotyšsku; správní středisko stejnojmenného kraje

Samostatné, číslované významy řadíme primárně podle kritéria frekvenčního, tedy podle pravděpodobnosti, že se s nimi uživatel v textu setká. Ostatní hlediska (viz např. Lew 2009, Adamska-Sałaciak 2006: 93n. aj.) jsou relevantní pouze v případě nedostatečné průkaznosti korpusových dat (což se však vzhledem k velikosti lotyšského korpusu stává nezřídka). Tak se na čelních místech – oproti jiným lotyšským slovníkům – často objeví druhotné, odvozené a přenesené významy.

**vandāllis** m<sub>2</sub> (~ja); **vanda||lis** m<sub>2</sub> (~ja) **1.** *vandal, chuligán* || *barbar, ničitel* || **2.** <sup>?</sup>pl ~li <hist> *Vandalové* (starý germánský kmen) || **3.** <IT> *malware*<sup>220</sup>

<sup>218</sup> V takovém případě nahrazují tradiční lexikografickou distinkci mezi čárkou a středníkem ve výčtech ekvivalentů. Je zřejmé, že hranice mezi samostatnými významy a pouhými nuancemi je těžko jednoznačně vytyčitelná a bude třeba rozhodovat se případ od případu, mj. též s přihlédnutím k typově blízkým heslům. Podstatné nicméně je, že samotná struktura databáze počítá s oběma eventualitami i s jednoduchým a pohodlným přepínáním mezi nimi (viz 2.4).

<sup>219</sup> Lépe než řešení, jež by mohlo mylně evokovat polysémnní povahu daného lexému:

**aitasgalvja** m/f<sub>4</sub> (Dsg m ~am, f ~ai) <hanl> **1.** *hlava skopová* || **2.** (muž) *hlupák, beran* || **3.** (žena) *hlupač-ka*

<sup>220</sup> V LVK na dotaz *vand[aā][lɪ].\** dostaneme 24 shod, z toho 14 s druhým vokálem dlouhým, 10 s krátkým. Ve 14 případech ze 24 jde o lexém *vandālisms/vandalisms* (zde je poměr 10:4 ve prospěch dublety

### 3.2.5 Informace syntagmatická

Syntagmatika, již jsme v kap. 2.2 přisoudili roli jednoho ze tří pilířů LČ, zahrnuje v našem pojetí tři různé jazykové jevy ve dvou rovinách: gramatické (valence) a lexikální (kolokabilita, představovaná kolokacemi a frazémy a také volnými kombinacemi uváděnými jako příklady). Těm odpovídají příslušné strukturní jednotky v naší databázi (atribut Valence a elementy Kolokace a Frazém – viz 2.4). Oprávněnost spojení těchto tří složek dokazuje následující příklad, kde jsou všechny tři kompatibilně začleněny do mikrostruktury hesla – společně vymezují základní kombinatorický potenciál lexému.

**asi|ns** f<sub>6</sub> (~ns, Gpl ~ŋu) *krev* • izliet ~nis (*par ko*) *prolīt krev* (*za koho/co*) • nodot ~nis *darovat krev* • ~ns lāse *kapka krve* • ~nīm piepludušas acis oči *podlītē krvī* • Kristus miesa un ~nis <náb> *tēlo a krev Pānē* [...] || <anat, med> (v odborném diskurzu převažuje singulár) • ~ns analīze *rozbor krve* • ~ns saindēšanās *otrava krve* • ~ns pārliešana *transfūze krve* • ~ns noņemšana *odběr krve* • ~s grupa *krevnī skupina* • ~ns aina *krevnī obraz* • arteriālās / venozās ~nis *arteriālnī (tepennā) / venóznī (žilnī) krev* • ~ns sastāvs *skladba krve* • ~ns vēzis *rakovina krve, leukēmie* [...] || ♦ sarkans kā ~nis *červený (rudý) jako krev* ♦ ~nis sakāpj (saskrien) *galvā krev stoupá do hlavy* ♦ ~nis sastingst dzīslās *krev tuhne v žilách* ♦ viņai dzīslās rit zilās ~nis *koluje jí v žilách modrá krev* ♦ gēnos viņam dažādu tautu ~nis *v genech má krev různých národů* [...]

3.2.5.1 Valenční údaje úzce souvisí s gramatickou charakteristikou lemmatu,<sup>221</sup> respektive jeho dílčích významů,<sup>222</sup> a zároveň komplementárně doplňuje syntagmatickou informaci nesenou jinými prostředky (glosami, exemplifikací). Není-li to na újmu přesnosti a přehlednosti, mohou být jimi i zastoupeny.<sup>223</sup>

---

s grafémem *ā* – i u tohoto lemmatu bude potřeba uvést obě varianty), všech zbývajících 10 výskytů (s převažující dubletou krátkou: doložena 6krát) se vztahuje k metonymickému významu ‚ničitel, barbar‘. Tyto ekvivalenty jsou také odděleny od dvojice *vandal*, *chuligān* jako významové nuance – do většiny kontextů uvedených v korpusu se nehodí, přesto je dobré je uživateli nabídnout.

Druhý význam, etymologicky původní, zaznamenáváme s příslušným diatechnickým markerem, převažující plurálový úzus je třeba ještě ověřit v jiných zdrojích. Rovněž je nezbytné prokázat oprávněnost zařazení třetího významu, který je na *tezaurs.lv* převzat z jediného terminologického slovníku, v ostatních zdrojích však doložen není.

<sup>221</sup> Srov. např. Čermákovo pojetí valence jako „diskrétní část[i] lemmatu“ (Čermák 1995d: 241).

<sup>222</sup> K paradigmatickému aspektu valence viz Čermák 1995c: 105.

<sup>223</sup> Tak může být syntagmatika, zvláště je-li odlišná od české, naznačena vhodným příkladem např. u zkratk:

**kdze**; **k-dze** (= kundze) *panī* (*p.*, *pī*) • *Liepiņas kdze p. Liepiņovā*  
**ist.** (= istaba) *pokoj, mīstnost* • *ist. 5 pokoj č. 5 • 5 ist. 5 pokojū*

Samotná valence pojímá „všechny u daného lexému systematicky predeterminované třídy a kategorie, signalizované obvykle přítomností některého formálního exponentu (pád, prepozice, konjunkce, adverbium, infinitiv, ale třeba i komparativ aj.), které se spolu s ním musejí, popř. mohou vyskytnout“ (Čermák 1995d: 242). Dřívější praxe omezovala valenci pouze na verbum jakožto syntaktické jádro věty, případně i na deverbativa – tuto obecnou schopnost lexémů je však nutno rozšířit i na další slovní druhy (ibid., Čermák 1995c: 105) a v LČ ji též náležitě reflektovat – a to jak u lotyšských lexémů, tak u jejich českých protějšků. Objevit se může nejen u lemmatu či dílčího významu, ale též u kolokace či frazému.

Formálně jsou údaje o valenci zapisovány podobně jako glosy (viz výše 3.2.4.1), pro snazší rozlišení však kurzivním písmem.

**māc[īt]**  $v_3$  (~u, ~i, ~a; ~tju) **1.** (*kam ko/infin*) *učit, vyučovat* (*koho co/čemu/infin*) • m. bērnu lasīt / svešvalodas *učit děti čist / cizí jazyky* • m. (*kādu*) darbā, m. (*kādam*) darbu *zaučovat (někoho) v práci [...]* || **2.** (*0 / ko kur*) (pracovat jako učitel) *učit* (*0 / koho/co kde*) • mana māsa ~a pirmajā klasē *moje sestra učí v první třídě (prvnířáky)* • viņa ir ~tjusi jau vairākas paaudzes *učila už několik generací [...]* || **3.** (*ko ko/infin*) *učit, vychovávat* (*koho co/k čemu/infin*), *vést* (*koho k čemu*) • man ~tja ar cieņu attiekties pret vecākiem cilvēkiem *učili mě chovat se uctivě (byl jsem veden k úctě) ke starším lidem* || **4.** (*ko ko/par ko/kā*) *poučovat* (*koho o čem/jak*) • ne~i mani, kā jāceļ māja *nepoučuj mě, jak [mám] postavit dům (jak se staví dům)* || **5.** (*ko*) (zvíře, přen. též člověka) *trénovat, cvičit, drezírovat* (*koho*) • m. suni *trénovat psa* || ♦ es tevi ~tšu! *já tě naučím! já ti ukážu!* ♦ ola grib vistu m. *kuře chce být moudřejší než slepice, kuře chce poučovat slepici [jak hrabat]*

**ņemot** partic: • ņ. vērā (*ko*) (ve funkci nevlastní předložky) *s ohledem na (koho/co), vzhledem k[e]* (*komu/ čemu*) též ↓ņemt [...]

**slā[pes]**  $f_5$  pl (Gpl ~pju) **1.** *žízeň* • just / sajust ~pes *mít / dostat žízeň* • [no]dzēsēt ([no]veldzēt) ~pes *hasit × uhasit, zahnat žízeň* • nomirt no ~pēm *umřít (zahynout) žízní [...]* || **2.** (*pēc kā*) (silná touha) *hlad, žízeň, lačnost (po čem)* • s. pēc demokrātijas / zināšanām *hlad po demokracii / vědomostech (znalostech)*

**gatav[s]** adj **1.** (*kam/infin*) *přípravený, přichystaný, nachystaný* (*k čemu/infin*) • g. lietošanai *přípravený k užívání* • g. piedalīties / uzpurēties *přípravený zúčastnit se / obětovat se* • valdība ir ~a pieņemt lēmumu *vláda je připravena přijmout rozhodnutí* • jābūt ~am dažādiem dzīves pavērsieniem *je potřeba být přichystán na různé životní zvraty [...]* || **2.** (*obv. součást přísudku*) *hotový, dodělaný, dokončený* • ēdiens ir jau g. *jídlo už je hotové* • drīz būs g. ekspertu *slēdziens brzy bude hotový znalecký posudek [...]* || **3.** (plody, zvl. ovoce) *zralý, uzrálý, dozrálý* • ~i āboli / tomāti *zralá jablka / rajčata* • kvieši jau ~i *pšenice už dozrála* || **4.** <hov> ~ais *hotový, učiněný, úplný* • ~ā elle! *hotový (učiněný) peklo!* • pianists ~ais (obrácený slovosled může signalizovat ironické užití) *hotový klavírista [...]* || **5.** <hov> *mrtvej, pošlej* • suns jau g. *pes už pošel* || ♦ nākt pie visa ~a *přijít k hotovému* ♦ dabūt ~u (*ko*) **a)** *poradit si (s čím)* **b)** *dokončit* (něco těžkého, nebezpečného ap.) [...]



**viņpus**<sup>1</sup> prep (+G/Dsg / D-Ipl) *za (čīm; co) || na druhé straně (čeho), na druhou stranu (čeho)*  
 • v. upes (upeī) / frontes (frontei) **a)** *za řekou / frontou (válečnou), na druhé straně řeky / fronty*  
**b)** *za řeku / frontu, na druhou stranu řeky / fronty* • dzīvot v. robežas (robežai) *žít za hranicemi, žít na druhé straně hranice* • v. okeānam to vērtē kā lielu veiksmi *za oceānem to hodnotí jako velký úspěch* • viņš arī redz v. notīm un atklāj muzikas struktūru *vidí i za noty a odhaluje strukturu hudby*<sup>224</sup>

3.2.5.2 Pod element Kolokace zaznamenáváme kromě samotných ustálených slovních spojení také vybrané volné kombinace lexémů (v heslové stati jsou jim kolokace většinou předřazeny).<sup>225</sup> Oba typy slouží primárně k ilustraci a konkretizaci typického úzu daného lemmatu, volná syntagmata – citovaná (případně parafrázovaná) vesměs z korpusu – mohou nadto signalizovat možné posuny a extenze významu. Spolu s frazémy tvoří příkladovou část, jejíž objem je odvozen od frekvence a/nebo sémantické komplexnosti hesla.

K vyhledávání kolokací (a ověřování aktuálnosti těch, jež uvádějí aktuální slovníky) používáme korpusové nástroje, především asociační (kolokační) míry. Lotyšské korpusy LVK a LVTK, dostupné přes korpusový manažer Bonito, pracují s mírami MI-score a T-score; české rozhraní KonText, v němž lze vyhledávat v rámci česko-lotyšské části paralelního korpusu IC, umožňuje využití i dalších měr: MI3, MI.log\_f, logDice a log-likelihood.<sup>226</sup> Orientovat se je možné rovněž podle bezprostředního okolí slova pomocí frekvenční distribuce. Cenné údaje přinášejí lexikální profily slov, generované nástrojem Word Sketches (viz 2.3.1.1).

Každý prvek množiny kolokací je pro přehlednost uvozen symbolem kolečka; dovoluje-li to analogie příkladů i jejich českých protějšků, využíváme zkrácené formy zápisu (znak lomítka označující alternativní lexém, v kulaté zá-

<sup>224</sup> Srov. minimalistická (a nápadně si podobná) řešení v jiných dvoujazyčných pasivních slovnících:  
**viņpus** *praep.* (z *gen. sg. lub dat.-instr. pl.*) *po tamtej stronie*; v. upes – *po tamtej stronie rzeki* (LP: 1467)  
**viņ** || [...]~**pus** *praep.* (*su gen sg. ir dat.-instr. pl.*) *anapus*; v. upes – *anapus upēs* (LL: 780)  
**viņpus** (*vsk. ar ģ, arī ar d., dsk. ar d.-i.*) *по ту сторону*; v. upes (upeī) *по ту сторону реки, на той (другой) стороне реки* (LR: II/680)

<sup>225</sup> Lze uvažovat i o alternativním řešení, kdy pro volná syntagmata a příklady zavedeme samostatný element a odlišíme je od kolokací i graficky (např. kolokace vysázet půltučně). Problematické je však jasné stanovení hranice, které syntagma je už ustálené a které ještě volné. Mimoto by další typ či řez písma mohl působit rušivě.

<sup>226</sup> [http://wiki.korpus.cz/doku.php/pojmy:asociacni\\_miry#asociacni\\_kolokacni\\_miry](http://wiki.korpus.cz/doku.php/pojmy:asociacni_miry#asociacni_kolokacni_miry).  
 O kolokačních mírách v češtině též Křen 2006.

vorce bývají synonymické prostředky). Řazení je většinou dáno frekvencí s přihlédnutím ke složitosti struktury daného příkladu.

**sakot** partic: • ar vienu vārdi s. *jedním slovem* • citiem vārdiem s. *jinými slovy [řečeno]* • paltiesību (taisnību) s. *popravdě řečeno* • godīgi (atklāti) s. *upřímně řečeno* • pareizāk s. *lépe (přesněji) řečeno* • īsāk s. *stručně řečeno, zkrátka, ve zkratce* • tā s. *řekněme, jak se říká, takřikajíc* || též ↑sacīt

**man|s** pron<sub>m1</sub> *můj* • m. tēvs / brālis *můj otec / bratr* • ~a māte / māsa *moje (má) matka / sestra* • ~s [vis]labākais draugs *můj nejlepší přítel (kamarád)* • pēc ~a prāta, pēc ~ām *domām podle mého názoru (mínění), podle mě* • tas ir m. *to je moje [...]*

**īsi** adv **1.** (prostorově) *krátce* • ī. apgriezti mati *nakrátko ostříhané vlasy* || **2.** (časově) *krátce* • ī. pirms (+Gsg, D-Ipl) *krátce, nedlouho před (kým/čím)* • ī. pēc (+Gsg, D-Ipl) *krátce, nedlouho, brzy, záhy po (kom/čem)* • tas notika ī. pirms Ziemassvētkiem *stalo se to krátce před Vánoci* • viņi atnāca ī. pēc astoņiem *přišli krátce po osmé* • ī. (~āk) sakot *stručně řečeno, zkrátka, ve zkratce* • ī. un skaidri *krátce (stručně) a jasně*

3.2.5.3 Výklad o syntagmatické informaci uzavře stručná zmínka o frazémech, jež budou pro svou specifickou zpracovány samostatně v pozdější fázi tvorby LČ. Klíčový je přitom jejich výběr, s ohledem na rozsah slovníku i jeho adresáta: právě lexémy idiomatické povahy dělají při překladu cizího textu největší potíže.

Frazémům je vyhrazen prostor v závěru heslové stati, od předešlé části jsou odděleny dvojitou svislicí a tradičně užívaným symbolem rombusu (ten je předsazen, podobně jako kolečko v případě kolokací, každé jednotce k snazšímu odlišení začátku a konce jednotlivých frazémů). Přestože tento přístup správně vyčleňuje frazeologii jako svébytnou složku lexikonu zvláště,<sup>227</sup> ani on není prost možných úskalí, popsanych F. Čermákem (1995a: 118) či B. Svensénem (2009: 194). Problematická je už samotná identifikace frazému: škála od volných slovních spojení přes ta méně či více ustálená (kolokace) až po frazémy nemá jednoznačné ohraničené předěly, záleží též na pojetí konkrétního lexikografa.

K lotyšským frazémům se snažíme najít jejich nejbližší významově-funkční český protějšek (s oporou především v SČFI a paralelním korpusu),

<sup>227</sup> Činí tak velká většina lotyšských slovníků. Z nám známých řadí frazémy pod jednotlivé významy pouze ČLN. Tento postup jde však proti základnímu rysu frazémů a tím je jejich sémantická nedělitelnost.

v nezbytných případech jej doplňujeme o stručnou vysvětlivku, raději než eventuální klasifikátor.

**lūgtin:** ♦ l. lūgt (naléhavě, intenzivně) *prosīt, žadonīt*

**ābeļle** f<sub>5</sub> (Gpl ~|u) *jabloň* [...] ♦ ābols no ~les tālu nekrīt *jablko nepadá daleko od stromu* ♦ kāda ā., tāds ābols *jaký pán, takový krám* ♦ balts kā ā. *bílý jako sníh* (vlasy, člověk)

**gatav|s** adj 1. (*kam/infin*) *přípravený, přichystaný, nachystaný* (*k čemu/infin*) [...] || ♦ nākt pie vi-sa ~a *přijít k hotovému* ♦ dabūt ~u (*ko*) **a**) *poradit si* (*s čím*) **b**) *dokončit* (něco těžkého, nebezpečného ap.) [...]

**proletārie|tis** m<sub>2</sub> (~ša) <sov> *proletář* || ♦ visu zemju ~ši, **savienojieties!** *proletáři všech zemí, spojte se!*

### 3.2.6 Informace diasystematická

Soustava diasystematických klasifikátorů, jimiž je LČ stejně jako ostatní překladové slovníky vybaven, přispívá k doplnění charakteristiky lexému po stránce sémantické (marker může pomoci desambiguovat konkrétní význam) či pragmatické (upozorňuje na expresivitu, evaluativnost a další rysy slova), především však poukazuje na jeho v jistém ohledu specifičnost, příznakovost, opozici vůči bezpříznakovému lexikonu, který zůstává neoznačkován. Oporou pro sestavení soupisu značek je nám koncepce diasystematického značkování F. J. Hausmanna (1989),<sup>228</sup> rozčleněná podle různých kritérií na 11 druhů. V této kapitole budou postupně analyzovány jednotlivé typy se zřetelem k LČ: některé z nich se do slovníku středního typu promítnou jen okrajově, jiné více, lze tu najít i jisté korelace.

Technicky je značkování v naší databázi zajištěno dvěma atributy různých elementů (Lemma, Význam, Kolokace a Frazém – viz 2.4): Stylový\_příznak a Oborový\_příznak,<sup>229</sup> přičemž seznam značek lze v TLexu podle potřeby snadno aktualizovat a doplňovat.

<sup>228</sup> Srov. též Hausmannův výrok: „Die Markierungssysteme der Wörterbücher stehen in der Regel auf schwachen theoretischen Füßen, sind aber dennoch unverzichtbar.“ (Hausmann 1985: 377, citováno podle Svensén 2009: 315). Diasystematické značkování z praktického hlediska popisuje např. Svensén 2009: 315–332,

<sup>229</sup> Atribut Oborový\_příznak vychází pouze z diatechnického aspektu (3.2.6.1), zatímco druhý atribut zahrnuje stylistické rysy, rozvrstvené mezi vícero hledisek popsaných níže. Pro lepší přehlednost databázového materiálu však toto zjednodušení tolerujeme.

3.2.6.1 Na základě časového hlediska vydělujeme značkování diachronické, stavějící proti současnému úzu jednak slova mrtvá, archaická, zastaralá či zastarávající, jednak neologismy. I přes koncepci LČ jakožto slovníku odrážejícího primárně současnou lotyštinu bude do něj v úzkém výběru zahrnuta i část ze starší vrstvy lotyšského lexikonu, už s ohledem na nejstarší generaci uživatelů jazyka. Tyto lexémy značíme markerem <zast>. Historismy chápeme jako historické termíny a opatřujeme je markerem <hist>, většinou jsou doplněny vysvětlující glosou.

**nevaids** (= nav) <zast> ↑būt

**kad<sup>2</sup>** konj 1. (uvozuje vedlejší větu časovou) *když* [...] || 4. <hov> ↑ka<sup>1</sup> || 5. <zast> ↓kaut<sup>1</sup> || [...]

**Abesīnija** f<sub>4</sub> <hist> *Habeš* (dnešní Etiopie)

**fogts** [o] m<sub>1</sub> <hist> *rychtář*, *fojt* (ve středověkém Livonsku a Německu úředník se soudní pravomocí)

Vzhledem ke stálé aktuálnosti sovětských reálií vydělujeme zvlášť specifickou skupinu dobově příznačných lexémů, které byly typické pro období sovětské okupace Lotyšska.

**biedr|s** m<sub>1</sub> (Vsg ~il) 1. (politické strany, organizace, klubu, komise ap.) *člen* [...] || 2. *kolega*, *druh*, *společník*, *kamarád* [...] || 3. <polit> *zástupce*, *náměstek* || 4. <sov> *soudruh* • b. Kalniņš *soudruh Kalniņš* • cienījamie ~i *vážení soudruzi* • ~u *tiesa soudružský soud*

**kriev|s** m<sub>1</sub> *Rus* [...] || ♦ iet / iesaukt ~os <hov, sov> *narukovat* / *povolat do armády*

**proletārie|tis** m<sub>2</sub> (~ša) <sov> *proletář* || ♦ visu zemju ~ši, savienojieties! *proletáři všech zemí, spojte se!*

Neologismy, běžně žádným klasifikátorem neoznačované, jsou obecně problematické; ani opakující se výskyt lexému dosud do databáze nezařazeného není sám o sobě zárukou: zrádné jsou módní okazionalismy a sezónní slova, jež se v jazyce nakonec neprosadí, přičemž jejich životnost se lexikografovi odhaduje jen těžko. Jejich případné přidání do hesláře by mělo být podmíněno dostatečnou mírou etablovanosti v jazyce, potvrzené z vícera relevantních zdrojů.

3.2.6.2 Od hlediska místního se odvíjí značkování diatopické, vymezující celonárodní jazykový úzus od místních nářečí. Ta jsou v lotyštině a idiolektech jejích mluvčích dodnes poměrně živá (Stafecka 2013: 253). Navzdory tomu nepočítáme s dialektismy do hesláře LČ, nejde-li o slova zároveň příslušející k jiné jazykové vrstvě. Stejně tak bude zanedbána exilová varieta lotyštiny, jíž mluví krajanská komunita v zahraničí (zvláště v USA, Austrálii, Irsku, Švédsku, Velké Británii aj.)

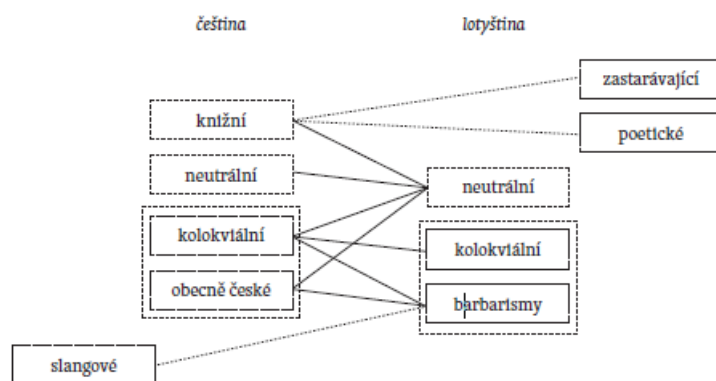
**dzirn|us** f<sub>3</sub> pl (Gpl ~u) <zast, dial> ↑dzirnavas

3.2.6.3. Hledisko národnostní zakládá značkování diaintegrativní, kladoucí do opozice slova domácí a zdomácnělá proti slovům cizího původu. V lotyšské lexikografii přetrvává praxe označovat speciálním markerem <barb> tzv. *barbarismy*, tedy negativně vnímané (a ze spisovné normy vykazované) prostředky přejaté z cizího jazyka či vytvořené podle cizího vzoru (Bušs 2005). LČ má být slovníkem deskriptivním, a tento normativní pohled mu tudíž nepřísluší (srov. též dále 3.2.6.11): nejčastěji dané lexémy, pocházející zejména z němčiny a ruštiny, nově i z angličtiny, klasifikujeme markerem <hov>, popř. <slang>, neboť jsou většinou typické pro mluvený jazyk.

**knifs** m<sub>1</sub> <hov> *finta, figl, grif*

**andelēt** v<sub>2</sub> (~ju, ~, ~; ~ju) <hov> (*ko*) *kšeftovat, handlovat (s čím)*

3.2.6.4 Médium v roli kritéria stojí za značkováním diamediálním. Neutrální lexikon je vymezen vůči prostředkům charakteristickým pro mluvený, či psaný jazyk. S. Nikuļcevová (2013: 219) konstatuje „nejsložitější a systémově nejodlišnější vztah“ mezi lotyštinou a češtinou právě v tomto ohledu, přehledně shrnujíc situaci do následujícího schématu, jímž se při hledání českých ekvivalentů řídíme i my.



**Obrázek č. 7:** Vzájemný vztah škály mluvenost–psanost mezi češtinou a lotyštinou (převzato z Nikuļceva 2013: 219)

Kolokvialismy bývají v normativních slovnících často zanedbávány (o manku vůči mluvené formě jazyka Čermák 1995c: 96, 1995d: 234), o to cennější je přístup aktuálního výkladového lotyšského slovníku MLVV, jenž v dosud nebývalé míře reflektuje i mluvenou lotyštinu (viz 2.3.2.1). Hovorově zabarvené lexémy opatřujeme markerem <hov> (či <slang>), a je-li to možné, tak i odpovídajícím obecněčeským ekvivalentem; stejně se dají označkovat ale i sublexikální prvky.

**šitād|s**  $\text{pron}_{m1}$  <hov> *takovej, takovejhle* • ~as muļķībās runāt! *takový kraviny říkat!*

**āzis**  $m_2$  (āža) **1.** [...] || ♦ dīrāt āzi <hov> *blejt, zvracet* [...]

**smēķis**  $m_2$  <hov> *cigáro, cigo*

**akme|ns**  $m_2$  (~ns / <hov> ~ņa, Npl ~ņi) *kámen* [...]

**psihene**  $f_5$  <slang> *cvokhaus, pakárna*

Naproti tomu pro prostředky psaného jazyka speciální marker nezavádíme;<sup>230</sup> jejich případný vyšší styl se snažíme postihnout ekvivalentem ze stejné stylistické roviny.

3.2.6.5 Hledisko sociokulturní vymezuje značkování diastratické, které od neutrálního lexikonu odlišuje sociolekty. Slangu, žargonu, popř. též argotu bude

<sup>230</sup> V lotyštině se knižnost vůbec nevyděljuje (Nikuļceva 2013: 219).

věnována pozornost až v pozdější fázi tvorby LČ, v omezeném množství by však zastoupeny být měly.

3.2.6.6 Hledisko formální definuje značkování diafatické, označující v opozici k neutrálním lexémům prostředky formální a neformální. Tato opozice má tradici spíše v anglosaském světě, v našem pojetí do značné míry splývá se značkováním diamediálním (viz výše 3.2.6.4), a nevyžaduje si tedy žádné specifické markery.

3.2.6.7 Od hlediska textově typového se odvíjí značkování diatextové, vydělující z neutrálního lexikonu různé typy slohů: prozaický, básnický a žurnalistický aj. Využíváme pouze druhý jmenovaný (marker <poet>), opět jen okrajově, spolu se zvláštním klasifikátorem pro slova z folklorních textů, zejména z lidové poezie, případně též pro různé lotyšské reálie (marker <folk>). Zavádíme rovněž speciální klasifikátor <bibl> pro biblismy. Odkaz na žurnalistický styl lze pak dobře kompenzovat glosou.

**dziesminieks** m<sub>1</sub> <poet> *pěvec, bard*

**guļns** f<sub>6</sub> (~ns, Gpl ~ņu) <poet> *oheň, plamen*

**kumeļš** m<sub>1</sub> **1.** *hříbě* **2.** <folk> *koníček, kůň*

**ačkups** m<sub>1</sub> <folk> lot. lidový tanec

**Ābels** m<sub>1</sub> <bibl> *Abel • Ā. un Kains Ábel a Kain*

**Lk** (= Lūkas evaņģēlijs) <bibl> *Lukášovo evangelium (Lk)*

3.2.6.8 Hledisko technické a od něj odvozené značkování diatechnické vyděluje z lexikonu odbornou terminologii, resp. různé oborové terminologie. Ty zpracováváme s oporou v terminologických slovnících a databázích (viz 2.3). Počet oborových markerů jde v některých slovnících do desítek až pár set; LČ volí střídmější přístup a zavádí značku pouze v případě dostatečného zastoupení

daného názvosloví v databázi (řadově alespoň desítky lexémů).<sup>231</sup> Stejně tak nečleníme obory dále na jednotlivé disciplíny, především s ohledem na uživatele, pro něž by mohly být neprůhledné zkratky <ichtyol>, <entom> a <ornit>, spojené v LČ pod společný marker <zool>; obdobně není vydělen marker <gram> od <lingv>, <geom> od <mat> apod. Diatechnické klasifikátory na rozdíl od všech ostatních typů pro snazší orientaci podtrháváme.

**kvarťar|s** m<sub>1</sub> <geol> *čtvrtohory, kvartér* • ~a nogulumi *čtvrtohorní sedimenty* • ~a periods *období čtvrtohor*

**bedřil|te** f<sub>5</sub> (Gpl ~šu) **1.** *důlek, jamka, dolíček* • golfa b. *golfová jamka* || **2.** (na tvářích, bradě) *dolíček, důlek* || **3.** <anat>: • locitavas b. *kloubní jamka* • zoba b. *zubní lůžko*

**ağent|s** m<sub>1</sub> **1.** (firmy) *agent, zástupce, jednatel* [...] || **2.** (tajné služby) *agent* [...] || **3.** (hybná síla, původce děje) *činitel, agens* || **4.** <chem, med> *čínidlo, agens* • ķīmiskie ~i *chemická čínidla* || **5.** <IT> *agent*

**adadžo** [a'dadžo:] adv <hud> *adagio* (klidně, volně)

3.2.6.9 Z hlediska frekvenčního plyne značkování diafrekvenční, lišící slova běžně se vyskytující od těch řídkých či vzácných. Tento typ markeru nevyužíváme: frekvenční kritérium je pro nás prvořadé (viz 3.1.1), a není-li pro to pádný důvod, slova nedoložená či doložená jen spoře nejsou do hesláře LČ zařazena. Méně časté varianty klademe za ty frekventovanější, ať už vertikálně (v rámci makrostruktury) nebo horizontálně (v heslové stati). Řidší význam nebo jeho potenciální rozšíření či rozšíření lze pak naznačit vhodným příkladem.

3.2.6.10 Z hlediska postojového lze vyvodit značkování diaevaluativní, vyčleňující lexémy podle přítomnosti, či absence pragmatické složky významu. Využíváme běžné a praxí osvědčené markery <hanl>, <expr>, <iron>, <žert>, <eufem> aj., často je však lze nahradit explicitněji podanou informací v glose.

<sup>231</sup> Takto je např. v ČLN (12–14) v seznamu zkratek užitých ve slovníku uvedena zkratka *včel.* pro včelařství, v samotném hesláři se však tento marker objeví pouze dvakrát (*hrozen, leták*), v ostatních případech (*buňka, dělnice, hnízdo, roj, rojit se*) si autorka vystačila s glosou (*včelí, popř. včely*). Podobně jsou v ČLN uměle zaváděny další oborové markery: *archeol.*[ogie] 3 výskyty, *film.*[ařství] 5, *horn.*[ictví] 5, *mysl.*[ivectví] 5, *námoř.*[nictví] 2, *obch.*[od] 2, *text.*[ilnictví] 1, *žel.*[ezniční obor] 2 a dalších 11 markerů s méně než 10 výskyty v celém slovníku.



**drank|is** m<sub>2</sub> <hov, hanl> **1.** (nekvalitní výrobek) *šmejd, brak, škvár* • ~a filma *béčkový (brakový) film* • dzert visāds ~us *pīt všelījāký patoky* • ~a laiks *počasī pod psa* || **2.** (člověk) *hajzl, mize-ra, šmejd*

**gatav|s** adj [...] **4.** <hov> ~ais *hotový, učiněný, úplný* • ~ā elle! *hotový (učiněný) peklo!* • pianists ~ais (obrácený slovosled může signalizovat ironické užití) *hotový klavírista* [...]

Zpracována by měla být rovněž třída vulgarismů, jež jsou pro mnohé lexicografy stále ještě tabu.<sup>232</sup> Je ale žádoucí explicitně uživatele upozornit na možné společenské sankce, použije-li daný výraz; činíme tak speciálně zvýrazněným markerem.

**dir|st** v<sub>1</sub> (~šu, ~s, ~š; ~su) <vulg!> **1.** *srát* || **2.** *kecat, žvanit* (vymýšlet si n. mluvit hlouposti) || ♦ ej d.! *jdi (běž) do prdele!* ♦ beidz d.! **a)** *přestaň si vymějšlet!* **b)** *přestaň kafrat!*

**peža** f<sub>4</sub> <vulg!> *píča*

3.2.6.11 Konečně poslední hledisko, normativní, definuje značkování dia-normativní, jež vyděluje výrazy „správné“ a „nesprávné“, tedy s ohledem na jejich vztah k aktuální jazykové normě. Problematika preskriptivního pohledu na lexikon byla zmíněna v bodě 3.2.6.3 v souvislosti s barbarismy. Vedle nich bývají v lotyšských výkladových slovnících (a občas i v překladových) označovány další třídy slov: v LLVV, LLV i MLVV jsou to tzv. *vienkāršrunas vārdi* („slova běžného [dosl. prostého] jazyka“), používaná „v nespisovné hovorové řeči a svědčící o nedbalé mluvě a ledabylém postoji k výslovnosti a volbě slov“ (VPSV: 446), *nevēlami vārdi* („slova nežádoucí“ [ve spisovném jazyce]); MLVV nadto zavádí též klasifikátor *sar. nelit.* pro slova hovorová nespisovná.<sup>233</sup> Všechny tyto markery v LČ slučujeme do společného markeru <hov> a o jeho zařazení do hesláře rozhoduje rozšířenost a frekvence výrazu; specifika týkající se jeho úzu se dají doplnit na konec hesla (viz dále 3.2.7).

<sup>232</sup> V tomto ohledu je třeba ocenit rozhodnutí tvůrců LP zahrnout do hesláře řadu vulgarismů. Pořád jako by však přetrvával jakýsi ostych z těchto slov. Obě níže uvedená příkladová hesla mají v LP kuriozní podobu, připomínající Obrátilova Kryptadia.

**dir|st** v. (I, ~šu, ~s, ~š; ~sa; ~īs [sic!]) *wulg. 1. fizjol.* *wypróżnić się / wypróżniać się; załatwiać się; 2. pogard., przen. ściemniać; kłamać; mówić nieprecyzyjnie (pokrętnie); kręcić; ♦ beidz d.! pogard. – przestań kłamać!*

**peža** s. (IV) *žarg., anat.* *pochwa* (LP: 401, 1043)

<sup>233</sup> V MLVV je v seznamu zkratk uvedeno celkem 103 položek, z toho 10 je stylových markerů, v LVV je to 9 ze 48, v LLVV 10 ze 130 (Jērane 2013: 73).

Diasystematické značkování, zvláště to stylistické, je snad vůbec nejsubjektivnějším složkou lexikografovy práce; není náhodou, že stejné lexémy bývají ve slovnících popsány v tomto ohledu různě (četné příklady z lotyšského prostředí uvádí Jērane 2013). Příslušnost lexému k tomu kterému registru má sklon měnit se poměrně rychle, čímž zastarává i slovníkový popis takto označeného výrazu. Doporučuje se proto omezit značkování na minimum<sup>234</sup> a ve finální fázi ověřit ještě jeho aktuálnost (Čermák 2010b: 569, srov. též návrh vůbec nadřadit nadhodnocovanému stylu pro jeho vágnost frekvenci – Čermák 1995c: 99–100). Platí též, že ekvivalent by „měl optimálně podchytit implicity maximum shodných potřebných informací sám o sobě“ (Čermák 1995d: 240), tedy bez glosy a klasifikátorů.

### 3.2.7 Informace uzuální

Explicitní popis úzu daného lexému prostřednictvím elementu Úzus (viz 2.4.10) používáme všude tam, kde to není možné učinit jinými, stručnějšími prostředky (glosy, stylistické klasifikátory). Využití tohoto prvku, v lotyšské lexikografické praxi neznámého,<sup>235</sup> má široký potenciál: kromě upozornění na specifickou pragmatickou složku daného lexému jím lze zprostředkovat frekvenční údaje (např. srovnání četnosti dublet či alolexů aj.), představit uzanci v určité oblasti, resp. naznačit rozdíly v českém a lotyšském prostředí, upozornit uživatele na potenciální úskalí např. v podobě tzv. falešných přátel, doporučit

---

<sup>234</sup> „Z pragmaticko-praktických příčin by v slovníku malo byť čo najmenej kvalifikátorov: slovník stráca aktuálnosť predovšetkým ‚vdďaka‘ kvalifikátorom, keď každá lexikálna jednotka má svoju – v podstate dobovú – značku. Kvalifikátory v slovníku by však mali objektívne odrážať pohyb v jazykových podsystemoch (pravopis, výslovnosť, morfológickú a slovotvornú variantnosť a i.), nielen registrovať daný stav. Zároveň však kvalifikátory vydávajú svedectvo aj o jazykovej kompetencii tvorcov slovníkov, sú v nich v rozličnom stupni obsiahnuté aj autorské subjektívne postoje a hodnotenia. Hlavný problém je teda v tom, ako prostredníctvom kvalifikátorov čo najviac zjednotiť a zobjektívizovať tieto hodnotiace postoje aj pre používateľov slovníkov.“ (Bosák 2008: 123–124)

<sup>235</sup> Jako první ji zčásti zavádí MLVV, omezuje se však na hodnocení daného lexému z hlediska jazykové normy a/nebo kultury, což je jistě vliv silné tradice preskriptivismu v lotyšské jazykovědě (Jansone–Veisbergs 2013: 440). Srov. naopak silnou tradici tzv. *usage notes* v anglosaském světě, reprezentovanou především slovníkem *The American Heritage Dictionary of the English Language*.

mu z víceroch možností jednu konkrétní apod. Výklad může být doplněn odkazem na jiné heslo ve slovníku nebo vně hesláře na gramatický přehled (kde bude daný jev soustavně vysvětlen), lze jej opatřit výstižnými příklady spolu s českým překladem či jinými dodatky. Na rozdíl od zbytku heslové statě tu není bezpodmínečně nutná unifikace, poznámky vyhovují vždy individuální potřebě (které by ovšemželo kategorizovat), a každá je tudíž osobitá.

Rozmanitost nabízených informací naznačují příklady níže.

**sāl|s**  $m_2$  (~ls, Dsg ~im, Npl ~li) /  $f_6$  (~ls, Dsg ~ij, Npl ~lis, Gpl ~lu) **1.** *sůl* • vārāmā (vārāmais) s. *ku-chyňská sůl* • ~ls raktuve *solný důl* • piebērt ~li [zupai] *přisolit, dosolit* [polévku] [...] **2.** <chem> *sůl* • skābie / bāziskie ~li (skābās / bāziskās ~lis) *kyselý / zásaditý soli* [...]

Úzus: V odborném diskursu (2. význam) převažuje mužská podoba slova, častý je plurál; v neoborném (1.) jsou zase běžné tvary jednotného čísla s ženskými koncovkami

**ņemot** partic: • ņ. vērā (ko) (ve funkci nevlastní předložky) s *ohledem na* (koho/co), *vzhledem k* [e] (komu/ čemu) též ↓ņemt

Úzus: Tvar ņemot se v LVK objevuje 743×, z toho 663× (tj. téměř v 90 %) ve spojení s lexémem vērā, které má povahu sekundární předložky (též → **Gramatika: s. 1060**). Lze jej přeložit však i opisně: ņemot vērā šogad plānoto cenu pieaugumu tabakai, inflācija nevarēja nepieaugt. „Vezmeme-li v potaz letošní plánovaný nárůst ceny tabáku, nemohla neporůst ani inflace.“

**vaja|dzēt**  $v_3$  (~g; ~dzēja) vyjadřuje nutnost, nezbytnost, potřebu ap. • man ~g naudas *potřebuju peníze* • ko tev ~g? *co potřebuješ?* • ko jums ~g? (v obchodě, na úřadě ap.) *co byste potřeboval?* [...] || *být potřeba, být zapotřebí* • ~g pieņemt lēmumu je *potřeba přijmout rozhodnutí* [...]

Úzus: Statistika tvarů v LVK: vajag (přít. čas) 904, vajadzētu (podmiňovací zp.) 785, vajadzēja (min. čas) 625, vajadzēs (bud. čas) 170, vajadzējis (přičestí minulé činné) 52, vajagot (relativ) 38, vajadzēt (infinitiv) 2, vajadzēšot (relativ bud. času) 2. – Infinitiv se prakticky nepoužívá, dominuje užití ve 3. osobě ve všech časech a kondicionál.

**mēr|s**<sup>2</sup> [e:]  $m_1$  *starosta*, (velkoměsta též) *primátor* [...]

Úzus: Starostu oslovujeme nikoliv prostým vokativem (mēr), ale spojením mēra kungs.

**čigān|s**  $m_1$  **1.** *Cikán, Rom* • ~u valoda *romština* [...]

Úzus: Na rozdíl od češtiny není v lotyštině slovo čigāns vnímáno jako pejorativní a používá se i v oficiálních situacích. Slovo roms je užíváno řidce (pouhé 4 výskyty v LVK proti 93 výskytům slova čigāns).

**lei|tis**  $m_2$  (~ša) *Litevec* • ~šu valoda *litevština*

Úzus: V určitých kontextech může být toto slovo považováno za hanlivé. Lépe je proto použít neutrálního lietuvietis, které je také daleko častější (v LVK 206× proti 18× leitis).

**stāv|s**<sup>2</sup>  $m_1$  **1.** *patro, poschodí, podlaží* • pirmais s. *přízemí* • otrais / trešais s. *první / druhé patro* [...]

Úzus: V Lotyšsku se poschodí počítají od přízemí. Trešais stāvs je tedy druhé, nikoliv třetí patro, podobně i trīsstāvu ēka je dvoupatrová budova.

**adre|se**  $f_5$  (Gpl ~šu) **1.** *adresa* • ~šu saraksts *adresář* • ~ses maiņa *změna adresy* • nosūtīt uz (kāda) ~si *poslat na (čí) adresu* [...]

Úzus: Adresa na poštovních zásilkách se objevuje nejčastěji v této podobě: Baibai Kalniņai (jméno adresáta v dativu), Burtnieku ielā 21 (název ulice v lokativu), Rīgā (název města v lokativu), LV-1050 (odpovídá čes. PSČ), Latvija (název státu v nominativu).

**Tallin|a** f<sub>4</sub> Tallinn (metropole Estonska)

Úzus: V českých textech doporučujeme psát v souladu s originální podobou jména dvě l a dvě n.

**Zev|s** m<sub>1</sub> <mytol> Zeus

Úzus: V lotyštině se jméno Zevs skloňuje pravidelně (Zeva, Zevam, Zevu, Zevā), na rozdíl od češtiny (Zeus, Dia, Diovi...).

**divpunk|te** f<sub>5</sub> (Gpl ~šu) <lingv> dvojtečka (..)

Úzus: V lotyšských textech se dvojtečka (dvě tečky psané za sebou, nikoliv nad sebou) používá k označení vypuštěného textu a odpovídá tak české trojtečce ve funkci výpustky. Např.: Divpunkte latviešu valodas rakstos ir jauns pieturzīme, kas .. iegūst arvien stabilāku lietojumu. „Dvojtečka v lotyšských článcích je novým interpunkčním znaménkem, jehož (...) užívání stále získává na stabilitě.

**aicin|āt** v<sub>3</sub> (~u, ~i, ~a; ~āju) 1. (ko kur) (na návštěvu, do divadla, do restaurace ap.) **zvāt** (koho kam) [...]

Úzus: Když lotyšsky řeknete es tevi aicinu, neznamená to ještě, že chcete za druhého platit. Můžete sice, ale nemusíte. Jestli chcete dát jasně najevo, že hodláte i zaplatit, použijete k pozvání výraz es tev uzsaucu.<sup>236</sup>

**sirdīgs** adj (~ša) (člověk) rozzlobený, rozhněvaný, zlostný || (zvíře též) vzteklý

Úzus: Neplést s ↓sirsniġs.<sup>237</sup>

### 3.2.8 Informace referenční

Důležitou roli z hlediska soudružnosti makrostruktury a megastruktury slovníku plní jednoduchý a intuitivní systém odkazů, které navigují uživatele po hesláři (a to směrem zpět, či dopředu, což přehledně rozlišuje dvojí šipka ↑ / ↓), anebo jej odkazují vně hesláře na gramatický přehled lotyštiny (pomocí šipky → a klíčového slova *Gramatika*). Referenční systém je v TLexu zajištěn pomocí interního elementu Reference (viz 2.4.9),<sup>238</sup> který umožňuje nejen přidání nových typů odkazů, ale též flexibilní kombinování těch stávajících.

**liec, liec**<sup>-1</sup> préz ↓likt<sup>1</sup>, likties<sup>1</sup>, likt<sup>2</sup>, likties<sup>2</sup>

**nevaids** (= nav) <zast> ↑būt

**sākot** partic: • s. ar (ko), s. no (kā) (ve funkci nevlastní předložky) *od* (koho/čeho), *počínaje* (kým/čím) • plānots, ka, s. no 2007 gada, fondā būs pieejami 8 miljoni latu *plānuje se, že od roku 2007*

<sup>236</sup> Příklad citován z Běťáková 2012: 104.

<sup>237</sup> „Srdečný, upřímný, milý“; oba tvary jsou odvozeny od substantiva *sirds* „srdce“, mají však výrazně odlišnou sémantickou prozodii.

<sup>238</sup> Informace z neveřejného atributu Ref (2.4.3.20), sloužícího interním potřebám sestavovatele databáze, se do LČ nepromítanou; jde tu především o budování referenční sítě na základech sémantických, případně slovtvorných. Vedle sémaziologického vystavěného překladového slovníku je tu tak potenciál i pro využití onomaziologického přístupu (tvorba tezauru, synonymických soupisů apod.).

bude ve fondu k dispozici 8 milionů latů || též ↓sākt

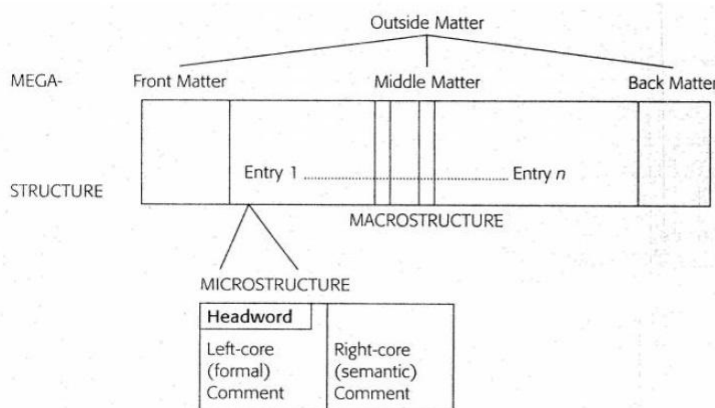
Úzus: V LVK se tvar sākot objevuje 466×, z toho ve spojení s předložkami ar 217× a no 149×. Téměř v 80 % případů tak plní funkci sekundární předložky (též → Gramatika: s. 1060).

### 3.3 Možnosti megastruktury

Právě popsaná dvojí základní struktura slovníku – makrostruktura a mikrostruktura – bývá v lexikografické teorii tradičně zastřešována tzv. megastrukturou, jež určuje pořadí a vztahy mezi jednotlivými složkami slovníku (Hausmann–Wiegand 1989: 330–333, Bergenholtz–Tarp 1995: 211–214, Svensén 2009: 379). Centrální částí je vždy samotný heslář, v LČ pouze jediný, zahrnující i typy hesel, jež bývají v jiných slovnících vyčleňovány do samostatných vedlejších heslářů (viz 3.1.2.5). Ostatní materiál se v závislosti na pozici vůči hesláři tradičně rozděluje na tzv. *front matter*, jež hesláři předchází, *back matter*, jež po něm následuje, a *middle matter*, který je do hesláře vložen, aniž by však byl jeho součástí (nejčastěji se jedná o obrazové přílohy).<sup>239</sup> S posledně jmenovanou složkou LČ nepočítá, heslář bude doplněn o následující informace:

V krátkém úvodu bude představen samotný slovník (po stránce metodologické i obsahové, včetně uvedení hlavních zdrojů – viz též 2.3) a stručně také lotyšština jakožto objektový (výchozí) jazyk spolu se specifiky jejích variet

<sup>239</sup> Všechny tři složky pak bývají souhrnně nazývány *outside matter*. Srov. graficky znázorněnou strukturu slovníku podle R. K. K. Hartmanna (2001: 59; viz též Hartmann – James 1998: 92).



Obrázek č. 8: Obrazové schéma struktury slovníku (převzato od R. R. K. Hartmanna)

(zejména té exilové, od níž LČ nicméně odhlíží, stejně jako namnoze od nářečních podob). Následovat budou oboujazyčné (ideálně v zrcadlovém podání) pokyny pro uživatele a seznam zkratk a symbolů ve slovníku užitých; jako zvláště názorný a edukativní se jeví schematický nákras hesla s legendou k jeho jednotlivým částem. Vzhledem k výsledkům četných výzkumů, podle nichž instrukce nechte drtivá většina uživatelů,<sup>240</sup> považujeme za nezbytné umístit tuto informaci ve zhuštěné podobě ještě na jiné, prominentní místo, kde s ní bude uživatel chtít nechtít trvale konfrontován.<sup>241</sup> Samotná struktura slovníku i meta-jazykový aparát by měly být co nejintuitivnější a nejjednodušší.

Po hesláři následuje stručný, hojně tabelizovaný výklad lotyšské gramatiky, na nějž se bude z hesláře přímo odkazovat (srov. ČLN 383–409). Vhodné je seznámit českého uživatele rovněž se zásadami transkripce a skloňování lotyšských slov, zejména proprií. Naopak lotyšský uživatel bude upozorněn na některá možná úskalí z jeho perspektivy (v hesláři nepostižené systémové dublety českých ekvivalentů apod.). Tyto přílohy budou pouze v jazyku cílového adresáta.

### 3.4 Editace a revize rukopisu LČ

Jen pro úplnost doplníme proces tvorby LČ i o fázi, která autora vznikajícího slovníku dosud čeká: fázi editační a kontrolní. I přes výše naznačené výhody TLexu (viz 2.4), jenž napomáhá konzistenci zpracování a předchází různým chybám a opomenutím lexikografovým, nelze tuto etapu zanedbat, je svým způsobem klíčová pro zdar celého projektu (srov. Balážová 2008). Po dokončení rukopisu je třeba nejprve revidovat nejasné a sporné případy (viz závěr kap. 4.3) a hledat řešení u analogických jevů. Stejně tak je nutno zkontrolovat homogen-

---

<sup>240</sup> Bo Svensén (2009: 459) s odkazem na H. Béjointa uvádí číslo až alarmující: 89 %. Srov. údaj 97 % (sic!) vzešlý z dotazníku mezi studenty filologických oborů Lotyšské univerzity (Veisbergs 1998: 155).

<sup>241</sup> Nabízí se např. předsádka knihy, jež bývají většinou beztak nevyužité: na přední by byl koncentrovaný návod pro české uživatele, na zadní týž pro uživatele lotyšské, v obou případech uvozený odkazem na kompletní výklad v úvodu slovníku.

nost popisu typově stejných hesel a tříd lexémů, mimo jiné též co se metajazyka týče, a eliminovat překlapy v obou jazycích (s využitím automatické kontroly pravopisu zabudované v textových editorech).<sup>242</sup>

Vedle samotných autorských revizí se plánuje rovněž využití rodilého mluvčího v roli redaktora lotyšské části LČ a konzultace s českými potenciálními uživateli slovníku.<sup>243</sup> V ideálním případě by měl LČ projít „zatěžkávací zkouškou“ v rámci zkušební, testovací fáze, kdy jej jak čeští, tak lotyšští uživatelé otestují na autentických textech různých typů. Relevantní připomínky a návrhy (včetně nových kandidátů na zařazení do hesláře) nekolidující zásadně s aktuální koncepcí by měl autor zohlednit a zapracovat do výsledné podoby slovníku.

Výhledově lze ovšemže uvažovat i o vzniku elektronické verze LČ, struktura databáze jí je uzpůsobena. Proti tištěné verzi by elektronický slovník, ideálně volně dostupný online, měl nesporné výhody v prakticky neomezeném prostoru (další nárůst hesláře o specifické složky lexikonu, omezení zkratk ve prospěch explicitního popisu, posílení bidirektoriálnosti apod.), v možnosti pravidelně aktualizovat heslář (přidáváním neologismů, resp. nových významů slov a naopak vyloučením slov z úzu mizejících) a opravovat chyby, nejasnosti a nekonzistence a v neposlední řadě též v pokročilejších a flexibilnějších technických možnostech (přidání zvukového a obrazového materiálu, robustnost fulltextového prohledávání<sup>244</sup>) a lepší funkcionalitě, což zajisté povede k většímu uživatelskému komfortu. Primární však i nadále zůstává tištěná podoba LČ, představující oproti případné dynamické elektronické verzi konkrétní synchronní řez a zůstávající tak dokumentem své doby.

---

<sup>242</sup> Srov. např. soubor opatření navrhovaný Honselaarem (2003: 330–331), viz též Čermák 1995d: 237.

<sup>243</sup> Opakované konzultace s budoucími českými i lotyšskými uživateli však probíhají již v průběhu tvorby slovníku.

<sup>244</sup> Srov. důvtipný výrok S. Atkinsové pronesený již v r. 1996: „At last we are liberated from the straitjacket of the printed page and alphabetical order.“ (Atkins 1996: 516).

## Část čtvrtá

### Vybrané otázky dvoujazyčné lexikografie a jejich řešení v LČ

V této části se snažíme na třech dílčích případových studiích demonstrovat komplexní charakter lexikografické práce: kombinuje se tu rovina teoretická s praktickou, analytická se syntetickou.<sup>245</sup> Po vytvoření definitivní podoby hesláře musí lexikograf opětovně, znovu a znovu, čelit výzvě, aby stávající lemma doplnil o zbytek heslové stati, či – nahlíženo strukturalisticky – aby k existujícímu tématu našel réma. To vše s vědomím spletené sítě vztahů syntagmatických i paradigmatických, ale též funkčních každého jednotlivého lexému (srov. Čermák 2010a: 256); žádný z nich neexistuje v jazyce izolovaně, což by měl dobrý slovník reflektovat.

V návaznosti na popis vybraných slovních druhů v kap. 3.2.3 se pro ilustraci zaměříme na dvě substantiva (*biedrs*, *ārs*) a tři konjunkce (*turpretī*, *kurpretī*, *iekams*). Vzorové lexémy byly zvoleny tak, aby ilustrovaly souběžně vícero komplexních jazykových jevů rozličné povahy a pokryly tak pokud možno všechny roviny jazyka. V prvním případě se dotýkáme problematiky polysémie (a řazení jednotlivých významů) a diasystematických markerů (3.2.6), v druhém narážíme na otázku lexikální variantnosti a specifičnost exemplifikace (potřeba širšího, i nadvětného kontextu) a posledním příkladem je ukázána problematika homonymie, respektive homografie (viz 3.2.3.5), omezeného lexikálního paradigmatu (srov. heslo *vajadzēt* v 3.2.3.2.3.2) a z toho plynoucích různých eventuálních podob výsledné heslové stati. Všem předkládaným přípa-

---

<sup>245</sup> Slovy Sue Atkinsové: „The craft of lexicography demands not only the ability to collect data, but also the ability to make sense of it. We have to be able to see the whole picture of how our headword behaves in natural language before we can decide what information is most needed by the people who will use the dictionary we are writing.“ (Atkins 2002: 25)



dům je pak společná snaha zjistit, nakolik odpovídá stávající lexikografický popis daných lexémů reálnému úzu, respektive vzorkům, které poskytují korpusová data (srov. rozdíly mezi tradičním a korpusovým přístupem – Čermák 2010b: 561n.).

#### 4.1 Zpracování polysémního substantiva *biedrs* v LČ

Polysémní charakter slova *biedrs*, zvláštní mimo jiné dobově specifickým významem (‘soudruh’), je zjevný při nahlédnutí do lotyšských výkladových slovníků i letmé analýze korpusových výskytů. MLVV rozděluje sémantiku lexému do tří významů, z nichž první je dále členěn do tří podvýznamů (dvojitě lomítko označuje podle legendy významovou nuanci):

**biedrs 1.** Partner při práci či studiu, s nímž existují věcně přátelské vztahy // Spoluúčastník nějakého dění či události, za nějakých okolností // Člověk mající blízké názory, podobné životní podmínky apod.; **2.** Člen (politické strany, společnosti apod.); **3.** V sovětském zřízení – oslovovací forma občanů, rovněž zdvořilostní forma, připojovaná před něčí jméno, příjmení či označení funkce

Autoři MLVV vycházeli zjevně ze staršího LLVV: druhý a třetí význam jsou shodné, definice prvního významu, poplatná své době, příliš akcentovala kolektiv (‘člen kolektivu [ve vztahu k ostatním členům tohoto kolektivu]’), a proto byla přepracována a sloučena se čtvrtým významem v LLVV (‘člověk něčím pevně spojený s jiným člověkem’).

Manuální analýza realizací lemmatu *biedrs* v LVK přinesla tyto výsledky: ze 776 výskytů odpovídá 1. významu 204 případů (26,3 %), 2. významu 497 případů (64 %) a 3. významu 25 případů (3,2 %). Zbývajících 50 výskytů (6,4 %) obsahuje kolokace typu:

*Republikas 8. Saeimas priekšsēdētājas <biedrs> Jānis Straume . Sēdes vadītājs .  
Labrīt* (,zástupce předsedkyně Saeimy‘ – místopředseda parlamentu)

*Reģistrācijas rezultāti Nolasa - Saeimas sekretāra <biedrs> A. Klementjevs Balsojumi*  
(,zástupce tajemníka Saeimy‘ – zástupce tajemníka parlamentu)

*reģistrēties! Lūdzu Saeimas sekretāra <biedru> nolasīt reģistrācijas rezultātus.  
12,5 % ierosināja Ministru prezidenta <biedrs> Ainārs Šlesers. Viņš norādīja,*  
(,zástupce ministerského předsedy‘ – místopředseda vlády)

*Ministru prezidenta un Ministru prezidenta <biedra> birojs ir Valsts kancelejas  
struktūrvienība*

*tautsaimniecības komisijas priekšsēdētāja <biedrs> Vents Armands Krauklis pauž  
Tautas partijas* (,zástupce předsedy národohospodářské komise‘ – místopředseda ná-  
rodohospodářské komise)

Z perspektivy cílového jazyka zakládá toto užití nový, samostatný význam, který není ve výkladovém slovníku postižen. Jde sice o význam omezený na politický diskurz (navíc se jedná, alespoň podle dokladů, čistě o lotyšské specifikum), nikoliv však okrajový, jak se lze přesvědčit na internetu.<sup>246</sup> Je to zástupné označení za lexém *vietnieks* (,zástupce, náměstek‘), aktualizované z meziválečného období,<sup>247</sup> které má zdůraznit nesubordinační, rovnoprávnou povahu vztahu mezi příslušnými osobami. Je tedy žádoucí zaznamenat tento význam do databáze mezi ostatní a jednotlivé významy přeřadit s ohledem na získanou statistiku jednotlivých užití a taktéž vymezit jejich typický úzus příslušným klasifikátorem.

**biedr|s** m<sub>1</sub> (Vsg ~i!) **1.** [člen] || **2.** [partner, kolega, druh, spolu–] || **3.** <polit> [zástupce, náměstek]  
|| **4.** <sov> [soudruh]

<sup>246</sup> Stenografické záznamy jednání poslanecké sněmovny jsou v LVK zastoupeny pouhými 2 % (viz též 2.3.1.1), ve speciálních korpusech *saeima* (viz 2.3.1.1) je četnost tohoto významu daleko vyšší – srov. seznam levostranných kolokátů lexému *biedrs* v korpusu *saeima-2.0* s níže uvedeným soupisem z LVK: *sekretāra* 1183, *partijas* 509, *priekšsēdētāja* 255, *prezidenta* 251, *sekretāres* 212. Internetové výskyty naznačují možnou extenzi užití i mimo oblast politiky (opakované užití ve statutech různých společností apod.), srov. též doklady *redaktora/redaktors biedrs/bedre* ,zástupce/zástupkyně redaktora/redaktorky‘ v LVTK (celkem 5 výskytů). Typický úzus tohoto významu však zůstává omezen na politickou sféru

<sup>247</sup> Dobový MEV (I/305–306) tento význam nicméně nezaznamenává. Potvrzeno též osobním sdělením prof. Jurise Baldunčikse, jemuž bych tímto rád poděkoval za nejednu konzultaci.

Podívejme se nyní na bezprostřední levostranné kolokáty lemmatu *biedrs* v LVK (syntaktika a interpunkce jsou označeny kurzívou; tučně jsou zvýrazněny případy relevantní pro „nový“, tj. třetí význam).

partija ‚(politická) strana‘ 74	<i>par 10</i> ‚o‘
birža 41 ‚burza‘	savienība 9 ‚svaz, unie‘
saruna ‚rozhovor‘ 39	grupa 9 ‚skupina‘
sabiedrība 37 ‚společnost‘ (krājaizdevu sabiedrība ‚úvěrové sdružení‘ 23; kooperatīvās sabiedrība ‚družstevní společnost, družstvo‘ 9)	asociācija 9 ‚asociace‘
komanda 31 ‚mužstvo, tým‘	klase 9 ‚třída‘
<b>sekretārs</b> 27 ‚tajemník‘	apvienība 9 ‚sdružení‘
arodbiedrība 21 ‚odbory‘	<i>ar 7</i> ‚s‘
savs 20 ‚svůj‘	kurss 7 ‚kurz‘
, 19	darbs 7 ‚práce, zaměstnání‘
organizācija 16 ‚organizace‘	istaba 6 ‚pokoj, místnost‘
<i>rv 14</i>	<i>no 6</i> ‚z, od‘
) 13	<i>tā 5</i> ‚tak‘
. 13	<i>būt 5</i> ‚být‘
klubs 13 ‚klub‘	<i>un 5</i> ‚a‘
cīņa 13 ‚boj‘	» 5
<b>prezidents</b> 13 ‚prezident‘ (ministru prezidents ‚předseda vlády, premiér‘)	studija 4 ‚studium‘
skola 11 ‚škola‘	" 4
	dienests 4 ‚služba‘
	<b>sekretāre</b> 4 ‚tajemnice‘
	<b>priekšsēdētājs</b> 4 ‚předseda‘

Kolokace uváděné ve výkladovém slovníku se v korpusu vesměs objevují, do LČ by bylo záhodno převzít kromě těch nejfrekventovanějších také ty, jež by mohly českému uživateli při překladu činit potíže. Zde už využijeme paralelní korpus IC. V něm se lexém *biedrs* objevuje celkem 1355krát, z toho nejčastěji (cca v 760 případech) je překládán jako *člen*,<sup>248</sup> což potvrzuje správnost výše

<sup>248</sup> V IC hledáme pomocí dotazu na slovní tvar *biedr.\** a paralelně k tomu v české části v rámci základního dotazu wpisujeme potenciální český ekvivalent. Hodnoty nejsou úplně přesné, neboť uvedený slovní tvar může

uvedené změny v řazení jednotlivých významů. Frekvenční distribuce pádových forem je shrnuta v následující tabulce.

Tvar	Fce	Nejčastější ekvivalenty v Trequ (fce ≥ 3)
v IC/Trequ		
<b>biedri</b> (Npl, Vsg)	424/171	<i>členy</i> 41, <i>členové</i> 35, <i>producentů</i> 17, <i>soudruhu</i> 16, <i>soudruzi</i> 13, <i>seskupení</i> 11, <i>členů</i> 8, <i>kamarádi</i> 5, <i>nečlenů</i> 4, <i>složeními</i> 3
<b>biedru</b> (A/Isg, Gpl)	330/127	<i>členů</i> 91, <i>členy</i> 16, <i>druha</i> 3, <i>členem</i> 3
<b>biedriem</b> (D/Ipl)	258/134	<i>členům</i> 49, <i>členy</i> 30, <i>členů</i> 10, <i>členové</i> 4, <i>kamarádům</i> 4, <i>producentům</i> 3
<b>biedrs</b> (Nsg)	192/68	<i>členem</i> 27, <i>soudruh</i> 16, <i>kamarád</i> 4, <i>člen</i> 3, <i>seskupení</i> 3
<b>biedrus</b> (Asg)	65/9	<i>členy</i> 3, <i>kamarády</i> 3
<b>biedra</b> (Gsg)	48/1	11 hapaxů
<b>biedram</b> (Dsg)	37/3	<i>kamarádovi</i> 4, <i>člen</i> 4, <i>rozděluje</i> 3
<b>biedros</b> (Lpl)	1/1	hapax <i>nebohé</i>
<b>biedrā</b> (Lsg)	0/0	

**Tabulka č. 1:** Tvary lemmatu *biedrs* v IC/Trequ<sup>249</sup> a jejich české ekvivalenty

S hledáním protějšků k mnohým kolokacím nebudou větší obtíže: v IC převažuje jedno či dvojí řešení, které přejímáme do databáze, přiřazující je k příslušným významům.

*partijas biedrs* = člen (politické) strany

*biržas biedrs* = člen burzy (cenných papírů)

zahrnout též nežádoucí lexémy (*biedre*, *biedrene*, *beidrība*, *biedroties* aj.), jak je však patrné z frekvenční distribuce, jde o odchylku v řádu jednotek. Dalšími ekvivalenty v IC jsou: *soudruh* 106, *kamarád* 74, *kolega* 48, *druh* 33, *přítel* 29, *spolužák* 21, *partner* 17, *společník* 16; zbytek (zhruba 250 případů, tj. necelou pětinu) tvoří kromě překladů opisných též nesprávně zarovnané věty.

<sup>249</sup> Srov. frekvenční distribuci tvarů v LVK: *biedru* 252, *biedri* 157, *biedrs* 129, *biedriem* 101, *biedra* 68, *biedram* 39, *biedrus* 27, *biedros* 2, *biedrā* 1 a LVTK: *biedru* 6187, *biedri* 4448, *biedrs* 3756, *biedriem* 3307, *biedra* 1638, *biedrus* 1107, *biedram* 740, *biedros* 56, *biedrā* 8; vyskytují se též dubletní tvary skloňované podle 2. deklinace: *biedris* (Nsg) 50, *biedrim* (Dsg) 6, *biedrī* 4.

*komandas biedrs* = spoluhráč, týmový kolega

*arodbiedrības biedrs* = člen odborů, odborář

*darba biedrs* = spolupracovník, kolega z práce

*klases / skolas / studiju biedrs* = spolužák, kolega z třídy / ze školy / ze studií

V případě vícero kandidátů na ekvivalent<sup>250</sup> je nutné omezit volbu na dva až tři nejčastější a/nebo nejpřirozenější v cílovém jazyce – lexikograf nemá suplovat práci překladatele. Převládají-li opisná řešení a žádný z ekvivalentů není ustálený, spatřujeme východisko v rozšíření kolokace o vhodný kontext, který následně přeložíme perifrasticky. Pestrou nabídkou překladatelských řešení (přičemž by jistě bylo možné najít i další, neméně správná) se vyznačuje např. velmi frekventované spojení *sarunu (sarunas) biedrs*, doložené v IC celkem 45krát.

*Šajās lappusēs Eiropa runā imperiālistu valodā nevis valodā, kas izrādītu cieņu **sarunu biedram**.*

Na těchto stránkách se mluví imperativním jazykem, nikoliv jazykem, který respektuje svůj **protějšek**.

*Finanšu tirgu un fiskālās politikas jomā mums ir lielisks **sarunu biedrs**.*

Máme vynikajícího **mluvčího** v oblasti finančních trhů a fiskální politiky.

*Tas izskatījās tā, it kā viņš sarunu biedram pastāvīgi stāstītu kādu pikantu noslēpumu.*

Vypadalo to, jako by neustále svěřoval **někomu** jakási pikantní tajemství.

*Es sapazīstināju brāli ar viņa pamāti, un ar Krīzentāla palīdzību **sarunu biedri** pārmija dažas frāzes.*

Seznámil jsem bratra s macechou a s Krízentálovou pomocí si [0] vyměnili několik frází.

*Lai nu kā būtu, Sančo bija uzticīgs **sarunu biedrs**, viņu sarunas nepārvērtās par tenisa spēli uz vienu tīkla pusi, kur dons Kihots tikai gremdēja bumbas, bet Sančo rāpās pāri sētai tās meklēt un atnest atpakaļ.*

Ať jakkoliv, Sancho byl věrný **partner** jejich **hovorů**. Jejich hovory nebyl tenis hraný jen na jedné půlce dvorce, kde Don Quijot střílel do sítě a Sancho jen běhal, hledal a nosil mu míčky.

*Eiropas iestāžu vidū Komiteja veic specifiskas funkcijas : tā visaugstākajā mērā ir organizētās pilsoniskās sabiedrības pārstāvības un diskusiju vieta, privileģēts **sarunu biedrs** šīs sabiedrības dialogā ar Savienības iestādēm.*

---

<sup>250</sup> Např. kolokace *cīņu (cīņas) biedrs* s 10 výskyty v IC je překládána tu jako *spolubojovník*, jindy jako *kamarád ve zbrani* či *druh v boji* (de facto doslovný překlad této kolokace) anebo volněji jako *spojenec*.

V rámci struktury evropských institucí plní výbor specifickou úlohu : je v pravém slova smyslu místem pro zastoupení a diskusi organizované občanské společnosti a privilegovaným **prostředníkem dialogu** mezi ní a institucemi Unie.

*Pret cilvēkiem viņš izturējās kā līdzīgs pret līdzīgu, tādēļ bija saistošs un patīkams sarunu biedrs.*

Byl proto milým a zábavným **společníkem**.

*Ļaujiet man dalīties divos piemēros no manas nesenās faktu noskaidrošanas misijas, kuri pierādīja, ka sarunu biedriem nav jāiet tālu, lai vadītu jēgpilnu dialogu.*

Dovolte mi, abych se s vámi podělila o dva zážitky z mých nedávných misí, při kterých jsem byla pověřena shromažďováním informací, ukazujících, že **účastníci rozmluvy** nemusí zacházet příliš daleko, aby vedli smysluplný dialog.

*Iededzis lampu, viņš paņēma podu un uzlika lampu uz tā, lai būtu gaišāks un labāk varētu saredzēt sarunu biedru.*

Když ji rozžehl, vzal hrnec a postavil lampu na ten hrnec, aby lépe svítila a aby bylo dobře vidět na **rozprávk**.

*Ir ārkārtīgi grūti sarunu biedra domu gājienu un apziņu.*

Je mimořádně náročné zasáhnout vědomí a mysl **osoby, s kterou právě hovoříte**.

Vhodným rozšířením kolokace např. na jednoduchou příkladovou větu (modifikovanou z LVK) *viņš ir lielisks sarunu biedrs* (dosl. „je to skvělý druh při rozhovoru“) získáme i minimálně dva přirozeně znějící české ekvivalenty *skvěle se s ním povídá* a *je to skvělý společník*; stejný příklad se dá navíc využít také jako kolokát lemmatu *saruna*.

Celé heslo po zapracování zmíněných poznatků pak vypadá takto:

**biedrīgs** m<sub>1</sub> (Vsg -i!) **1.** (politické strany, organizace, klubu, komise ap.) **člen** • sociāldemokrātu partijas b. **člen sociální demokracie** • arodbiedrības b. **člen odborů, odborář** • kultūras komisijas b. **člen kulturní komise** • ~a karte **členská legitimace** • ~a nauda **členský poplatek** • ~u sapulce **členská schůze** • kļūt par goda ~u **stát se čestným členem** • biržas b. <ekon> **člen burzy** (cenných papírů) || **2.** **kolega, druh, společník, kamarád** • darba b. **spolupracovník, kolega z práce** • skolas / klases / studiju b. **spolužák (kolega) ze školy / z třídy / ze studií** • dzīves b. **životní partner, druh** • komandas b. **spoluhráč, týmový kolega** • viņš ir lielisks sarunu b. **skvěle se s ním povídá, je to skvělý společník** • cīņu b. **spolubojovník, kamarád ve zbrani** • domu b. **ideový soupevník** • ceļa b. **spolucestující, soupevník** • istabas b. **spolubydlící** || **3.** <polit> **zástupce, náměstek** • ministru prezidenta b. **místopředseda vlády, vicepremiér** • Saeimas priekšsēdētāja b. **místopředseda parlamentu** || **4.** <sov> **soudruh** • b. Kalniņš **soudruh Kalniņš** • cienījamie ~i **vážení soudruzi** • ~u tiesa **soudružský soud**

Pro srovnání uvádíme též heslo zpracované v LP:

**biedr|s,~e s. (I, V) 1. |organizācijas loceklis|**  
człon|ek,-kini; kļūt par kluba ~u – zostać  
członkiem klubu; apvienības goda b. – członek  
honorowy stowarzyszenia; partijas b. – członek  
partii; ~a karte – legitymacja członkowska;  
2. |draugs, paziņa| towarzysz|,-ka; kole|ga,  
-žanka; skolas b. – kolega szkolny; sola b. –  
kolega z ławki; cīņu b. – towarzysz broni; ceļa  
~i – towarzyszy podróży

Obrázek č. 9: Zpracování lexému *biedrs* v LP

LP zahrnuje ženské odvozeniny pod společné lemma, LČ z důvodů uvedených v kap. 3.2.3.1.1.3 upřednostňuje rozdělení do samostatných hesel. Členění na jednotlivé významy zůstává stejné, úsporněji je však řešena příkladová část, uvádějící jen kolokace opakovaně se vyskytující v korpusech, eventuálně takové, které by mohly českého uživatele zmást.<sup>251</sup>

Sémanticky blízký lexém *biedrene* je zpracován opět zvlášť, neboť s lexémem *biedre* sdílí pouze dva ze čtyř významů; o synonyma nemůže jít ani s ohledem na dosti rozdílné kolokační profily: z levostranných kolokátů mají oba výrazy společné jen některé, zbytek je komplementárně rozdělen, jak vyplývá z LVTK.<sup>252</sup>

- levostranné kolokáty typické pro lexém *biedre* (horní index před kolokátem odkazuje k příslušnému významu): <sup>1</sup>*apvienības* (10), <sup>1</sup>*asociācijas* (61), <sup>1</sup>*biedrības* (14), <sup>1</sup>*goda* (12), <sup>1</sup>*kolēģijas* (6), <sup>1</sup>*komisijas* (19), <sup>1</sup>*organizācijas* (8), <sup>3</sup>*priekšsēdētāja/priekšsēdētājas* (10), <sup>1</sup>*PSKP [KSSS‘]* (9), <sup>1</sup>*savienības* (59);

<sup>251</sup> Např. je-li v lotyštině shoda v jmenném rodu, zatímco v češtině se používá generické maskulinum. Viz segment z IC (kolekce *Acquis*): c ) " *ražotājs* " *ir jebkura fiziska vai juridiska persona* [fem.], *kas ir biedre ražotāju organizācijā un kas šai organizācijai piegādā savu produkciju , lai to realizētu , ievērojot Regulā ( EK ) Nr . 2200 / 96 iekļautos nosacījumus – c ) " producentem "*  každá fyzická či právnická **osoba** , která je **členem** organizace producentů a dodává své produkty této organizaci za účelem uvedení na trh za podmínek stanovených nařízením ( ES ) č . 2200 / 96 .

<sup>252</sup> Zde je lexém *biedre* doložen přibližně (tvar *Asg biedri* homonymní s *Asg/Npl* lemmatu *biedris* zanedbáváme) 528krát, lexém *biedrene* 261krát, srov. poměr 31:21 v LVK.

- pro lexém *biedrene*: <sup>2</sup>*darba* (13, spojení *d. biedre* jen 2), <sup>2</sup>*domu* (4), <sup>2</sup>*istabas/istabiņas* (20), <sup>2</sup>*klases* (22, spojení *k. biedre* jen 2), <sup>2</sup>*komandas* (16, spojení *k. biedre* jen 1), <sup>2</sup>*kursa* (25), <sup>2</sup>*skolas* (23, spojení *skolas biedre* jen 1), <sup>2</sup>*sola* (4), <sup>2</sup>*studiju* (9);
- pro oba lexémy: <sup>2</sup>*ceļa* (*biedre* 4 / *biedrene* 5), <sup>2</sup>*cīņu* (5/2), <sup>2</sup>*dzīves* (19/13), <sup>2+1</sup>*grupas* (2/3), <sup>1</sup>*kluba* (19/5), <sup>2</sup>*sarunu/sarunas* (14/29).

Tyto poznatky se přirozeně odrazí na podobě obou hesel.

**biedr|e** f<sub>5</sub> 1. (politické strany, organizace, klubu, komise ap.) *členka* • Latvija ir šīs organizācijas b. kopš 2002. gada *Lotyšsko je členem této organizace od roku 2002* || 2. *kolegyně, kamarádka* • *dzīves* b. *životní partnerka, družka* || 3. <polit>: *zástupkyně, náměstkyně* • ministru prezidenta b. *místopředsedkyně vlády, vicepremiérka* • Saeimas priekšsēdētāja b. *místopředsedkyně parlamentu* || 4. <sov> *soudružka*

**biedre|ne** f<sub>5</sub> (Vsg ~n!, Gpl ~ņu) 1. *kolegyně, kamarádka* • *skolas* b. *spolužačka* • *istabas* b. *spolubydliči* • *darba* b. *spolupracovnice, kolegyně z práce* || 2. <sov> *soudružka*

## 4.2 Varianty konjunkcí *turpretī* a *iekams* a jejich zpracování v LČ

4.2.1 V lotyšských výkladových slovnících je konjunkce *turpretī* zpracována pod společným lemmatem s dubletou *turpretim* (MLVV, LLVV), srov. též jejich popis v LVG (658–659), kde jsou ve výčtu odporovacích spojek uváděny s lomítkem, viz též LVPPV: 817. Řazení odpovídá i frekvenčnímu poměru obou variant v lotyšských korpusech. V LVK nalezneme tvar *turpretī* 143krát, variantu *turpretim* 63krát, ve stomilionovém LVTK je situace mnohem pestřejší:

*turpretī* 1534

*turpretim* 763

*turpretīm* 72

*turpret* 12

*turpreti* 9

*turpretiem* 1

*turpretis* 1



Třebaže je kvůli problematické kvalitě textů v LVTK nutno tyto výsledky korigovat,<sup>253</sup> je oprávněné připustit existenci dalších variant kromě dvou nejčastějších, kodifikovaných: hybridního tvaru *turpretīm*, dále formy *turpreti* a okazionálních dublet *turpretiem* a *turpretis*. Poslední tři jmenované podoby jsou marginální a v LČ se neobjeví, přesto je záhodno zaznamenat je i s četností do databáze jako symptom možného posunu úzu, který u synonymního lexému *kurpretī* / *kurpretim* pozorovat nelze.<sup>254</sup>

Při hledání českých ekvivalentů využijeme paralelního korpusu IC. V něm je poměr základních variant takovýto: *turpretī* 571, *turpretim* 374, další podoby zastoupeny nejsou. Nejčastějšími ekvivalenty jsou: *zatímco* (339), *naopak* (140), *naproti tomu* (107) a *kdežto* (30).<sup>255</sup>

Příznačně je nutno pojmut exemplifikační část hesla. Typický úzus konjunkcí se nejlépe vyjeví na širším kontextu, často i nadvětném. Příklady tak budou tihnout k idiosynkratickým rysům, o to těžší je tu lexikografova volba.

**turpretī; turpretim; turpretīm** <hov> [vše -'pre-] konj *zatímco*, *kdežto*, *naopak*, *naproti tomu*, *na druhou stranu*, *zato* • zelta cena pieaugš, t. naftas cena kritīsies *cena zlata vzroste*, *zatímco cena ropy spadne* • man t. šķiet, ka.. *mně se naopak zdá, že...* • Izlikties par muļķi dažreiz ir ļoti gudri. T. izlikties par gudru vienmēr ir muļķīgi. *Vydávat se za hlupáka je často velmi chytré. Zato vydávat se za chytráka je vždy hloupé.*

Obdobně je zpracován i synonymní lexém *kurpretī* (s variantou *kurpretim*), přičemž korpusová data ukazují na minimálně dvojí možnou významovou nuanci: jednak najdeme v IC pro tvary *kurpretī* / *kurpretim* mnohem méně českých ekvivalentů (dominuje *zatímco*: 34 případů ze 42), jednak tyto tvary až na zanedbatelné výjimky neuvozují samostatnou větu (viz výše příklad *Izlikties par*

<sup>253</sup> Protože jsou součástí LVTK hojně i texty bez diakritiky, mohl by tvar *turpretim* teoreticky být i bezdiakritickým zápisem varianty *turpretīm*, podobně i tvar *turpreti* s 9 výskyty může být bezdiakritickou podobou varianty *turpretī*, ve 4 případech se však objevuje zjevně v textech s diakritikou. Zato tvar *turpret* je ve všech 12 případech výsledkem defektního zpracování textu.

<sup>254</sup> LVK: *kurpretim* 3, *kurpretī* 1; LVTK: *kurpretī* 30, *kurpretim* 14, *kurpretīm* 1; IC *kurpretim* 37, *kurpretī* 5.

<sup>255</sup> Srov. též výsledky v aplikaci Treq: pro *turpretī* *zatímco* 138, *přičemž* 6, *namísto* 5, *kdežto* 5, *naopak* 6, *naproti* 2, 5 hapaxů; pro *turpretim* *zatímco* 134, *naopak* 7, *naproti* 4, *avšak* 2, 4 hapaxy.

*muļķi...*), fungují pouze v rámci souvětí. Tyto poznatky zohledňuje i finální podoba hesla.

**kurpretī; kurpretim** [-'pre-] konj *zatīmco, kdežto* • *vīrieši ir galvenokārt pārstāvēti augstos amatos, k. sievietes pārsvarā ieņem zemus amatus* *muži jsou zastoupeni především na vysokých pozicích, zatímco ženy zastávají pozice převážně nízké*

Srovnajme zpracování obou lexémů v jiných překladových slovnících: LP (731–732, 1382) je nekonzistentní, v jednom případě jsou dublety traktovány pod společným lemmatem (navíc bez označení přízvuku na druhé slabice), v druhém pod dvěma různými hesly, z nichž jedno je odkazové. V LL (360, 737) jsou varianty začleněny pod společné lemma, na konci obou hesel je dodatečný odkaz na synonymum (**kurpret**||**ī, ~im** [...] viz *tēž turpretī*). LR (577/I, 606/II) volí minimalistické řešení: obě hesla jsou vybavena pouze dvěma ekvivalenty, jeden z nich je společný; chybí jakákoliv exemplifikace, která tak výrazně omezuje uživatelovu představu o chování lexémů. – Žádný ze slovníků neuvádí hybridní podobu *turpretīm*, očividně pro její nekodifikovanost.

4.2.2 Podřadicí časová spojka *iekams* má dokonce čtyři kodifikované varianty: *iekams, iekāms, iekam* a *iekām* (LVG: 662–663; LVPPV: 279). LLVV shrnuje všechny pod společné lemma *iekams*, LVV (353) zpracovává první s druhou pod jedním a třetí se čtvrtou pod druhým heslem (definice je stejná, pouze lemma *iekam/iekām* je vybaveno exemplifikací), MLVV uvádí pouze první variantu (další tři nejsou uvedeny ani jako její dublety). Popis v gramatikách či slovnících však postrádá jakoukoliv zmínku o frekvenci jednotlivých forem, tuto informaci shrnuje následující tabulka.

	LVK	LVTK	IC (Treq <sup>256</sup> )
<i>iekams</i>	23	260	252
<i>iekāms</i>	0	65	1
<i>iekam</i>	0	11	0
<i>iekām</i>	0	13	2

**Tabulka č. 2:** Varianty spojky *iekams* v korpusech lotyštiny

Při analýze korpusových dat vycházíme ze sémantické odlišnosti dvou základních českých ekvivalentů (*než* × *dokud* [*ne*–]), které konstituují samostatné významy (srov. totéž vydělení v LP, LL i LR), k nim hledáme další možné české protějšky. Kratší varianty bez koncového –s hodnotíme pro jejich nízkou četnost jako druhotné a kvůli možné kolizi v abecedním řazení s heslem *iekampt* je zpracováváme jako samostatné, odkazové heslo. Příslušný fragment v LČ pak vypadá takto:

**ieka|lt** v<sub>1</sub> [...]

**iekam; iekām** konj ↓iekams

**iekam|pt** v<sub>1</sub> [...]

**iekams; iekāms** konj **1.** *než, dříve než, předtím než* • es ilgi domāju, ie. piekritu *douho jsem přemýšlel, než jsem souhlasil* • ie. es beigšu, es gribētu pateikt.. *dříve než skončím, rád bych řekl...* || **2.** (se záporným tvarem slovesa) *dokud* • ie. nav par vēlu *a) dokud není pozdě b) [dříve] než bude pozdě* • ikviens uzskatāms par nevainīgu, ie. viņa vaina nav atzīta *saskaņā ar likumu každý je považován za nevinného, dokud není jeho vina uznána v souladu se zákonem* • finansējuma problēmas jautājumā nekas nemainīsies, ie. uzņēmējiem nebūs pieejami lēti kredīti *v otázce financování se nic nezmění, dokud podnikatelé nebudou mít k dispozici levné úvěry*

**iekāpšan|a** f<sub>4</sub> [...]

<sup>256</sup> Treq nabízí ekvivalenty pouze k variantě *iekams*, a to: *než* 21, *dokud* 17, *předtím* 5, *přehledu* 2 a dalších 9 ekvivalentů s jediným výskytem, s výjimkou *dříve* irelevantních.

### 4.3 Zpracování homonymního lexému *ārs* v LČ

Lexikografické příručky vykládají lexém *ārs* jako homonymní; obě formy se od sebe liší intonačně: první (s intonací taženou: [ãrs]) označuje plošnou míru (č. ‚ar‘) a je monosémní, druhá (s intonací lomenou: [ârs]<sup>257</sup>) má v MLVV i LLVV čtyři významy, jež spojuje sémantický rys vnějšího, venkovního prostoru. V obou případech jde o maskulina náležící k první deklinační třídě, v druhém navíc o singulare tantum. Tento výchozí stav zaznamenáváme do TLexu.

**ārs**<sup>1</sup> [ã] m<sub>1</sub> ar (100m<sup>2</sup>)

**ārs**<sup>2</sup> [â] m<sub>1</sub> sg [vnějšek, venek]

Následuje pohled do korpusů: na základě LVTK lze uvažovat ještě o podmíněném zařazení zkratky *ĀRS* (v korpusu najdeme též její dešifraci); všech 24 výskytů však pochází z jediného zdroje a nemá oporu v jiných pramenech, s výjimkou nečetných případů na internetu (zkratkové heslo tak může zůstat pouze v databázi, do hesláře LČ nebude pravděpodobně zahrnuto – jde nejspíš o ad hoc vytvořenou, neustálenou zkratku).

?**ĀRS** [a: er es] (= ātrās reaġēšanas spēki) ?jednotky rychlého nasazení (dosl. ‚síly rychlé reakce‘)

V lotyšských korpusech je v drtivé většině zastoupen druhý lexém. V LVK na dotaz [lemma="ārs"] dostaneme 165 výskytů, s následující frekvenční distribucí pádových forem:<sup>258</sup>

*āra* (Gsg) 64

*ārā* (Lsg) 61

*āru* (Asg, Gpl) 33

*ārs* (Nsg) 2

<sup>257</sup> Srov. též LVPPV: 88 (kde je pořadí významů nicméně opačné) či Grīsle 2008: 19; zde je uvedena starší podoba *ars*, jež je homonymní s tvarem 3. osoby indikativu futura slovesa *art* ‚orat‘.

<sup>258</sup> Tvary s velkým počátečním písmenem připočítáváme k příslušné minuskové formě, pokud nejde o homonymní tvar s vlastním jménem *Āris*.

Vzhledem k problematické lemmatizaci LVK ověřujeme ještě výskyt jednotlivých forem. Genitivní tvar *āra* má o 2 výskyty více (jde však o homonymní formu Gsg propria *Āris*), počty tvarů *āru* a *ārs* jsou stejné, radikálně však narostl počet výskytů tvaru *ārā*: ten se v korpusu objevuje celkem 804krát, z čehož v 743 případech je označován jako adverbium se stejným lemmatem, ve zbylých 61 případech jako tvar Lsg lemmatu *ārs* (nicméně pouze v jediném případě jde o jednoznačné substantivum, zbytek je tedy označován chybně).<sup>259</sup> Tomu odpovídá i chápání této formy lotyšskou jazykovědou: jedná se o ustrnulou pádovou formu, jež ze synchronního hlediska tvoří samostatný lexém adverbialní povahy (č. ‚ven; venku‘), ve slovnících výkladových i ve většině překladových bývá tradičně vyčleňována pod zvláštní lemma. Je ji tedy potřeba vydělit z jmenného paradigmatu a zpracovávat zvlášť (viz dále).

Hypotetický neadverbální význam Lsg (ať už ‚v aru‘ či jakési ‚ve vnějšku‘) lze pro potřeby LČ zanedbat. Dvojitý výskyt nominativního tvaru *ārs* je okázionální (v jednom případě jde o beletristický text),<sup>260</sup> dativní tvar *āram* není doložen vůbec, v LVTK jen jednou. Analyzujeme-li výskyty zbylých dvou tvarů, především co se týče jejich kolokability, ukazují se další omezení úzu: z 33 výskytů akuzativního tvaru *āru* je celých 31 ve spojení s předložkou (*uz* 29, *pa* 2) a má výrazně adverbialní povahu; ustálenost těchto spojení lze ověřit i v dalších korpusech (LVTK: *āru* 436, *uz āru* 382, *pa āru* 33 + bezdiakritické varianty: *aaru* 51 *uz aaru* 43, *pa aaru* 6; IC viz dále). Genitivní tvar *āra* se běžně objevuje ve dvojí realizaci: buď ve spojení s předložkou *no* (16 případů ze 66, srov. v LVTK 35 případů ze 795), které je antonymem víceslovného adverbialního výrazu *uz āru* (‚zvenčí‘ × ‚ven‘), anebo ve funkci přívlastku rozvíjí ná-

<sup>259</sup> Srov. situaci v LVTK, kde je všech 7094 výskytů tvaru *ārā* otagováno jako adverbium.

<sup>260</sup> *Dambim jālūst, un daudz ūdeņiem jāasniedz ārs.* ‚Hráz se musí protrhnout a mnoho vody musí vytéct ven.‘ (dosl. ‚...dosáhnout venek‘) *Aiz sienas nebūs ārs, bez gan saimniecības telpa.* ‚Za zdí nebude prázdné (dosl. venek‘), ale hospodářské prostory.‘ V takovýchto případech se však neobejdeme bez volnějšího překladu.

sledující jméno<sup>261</sup> a odpovídá tak českému relačnímu adjektivu ‚venkovní, vnější‘.

Posledně jmenované užití umožňuje interpretovat daný tvar jako ustrnulou genitivní formu přebírající adjektivní funkci; ve výkladových slovnících je uváděno jako samostatný význam, v pořadí druhý. Plurálové tvary, odkazující k lexému *ārs*<sup>1</sup>, nejsou v LVK doloženy vůbec, v LVTK jen sporadicky, což však může být dáno velikosti, případně složením korpusu.

Výskyty v paralelním korpusu IC shrnuje následující tabulka.

Tvar	Fce v IC/Trequ	Ekvivalenty v Trequ
( <i>āra</i> )	844/175	<i>ven</i> 97, <i>venku</i> 54, po 2 výskytech <i>vypadněte, pryč, výběhu</i> , 18 hapaxů
<i>āra</i> (Gsg)	67/27	<i>venkovních</i> 8, po 4 výskytech <i>vnějšího, zvenčí, venkovním, venkovní</i> , 3 hapaxy
<i>āru</i> (A/Isg, Gpl)	48/11	<i>vyčnívají</i> 3, <i>arů</i> 2, 6 hapaxů
<i>āriem</i> (D/Ipl)	10/8	<i>arů</i> 8
<i>āros</i> (Lpl)	9/8	<i>arech</i> 8
<i>āri</i> (Npl)	4/2	hapaxy <i>dotkl, arů</i>
<i>āram</i> (Dsg)	4/0 <sup>262</sup>	
<i>ārus</i> (Apl)	3/3	<i>arů</i> 3

**Tabulka č. 3:** Tvary lemmatu *ārs* v IC/Trequ a jejich české ekvivalenty

Data z IC porovnejme co do kolokability s výše uvedenými výsledky z lotyšských korpusů:

- *āru* (48 výskytů): *uz āru* 42 (opakující se ekvivalenty: *vně* 10, *ven* 5, *směrem ven* 4, *navenek* 2, *naruby* 2; spojení *izvirzīt uz āru* celkem 9krát pře-

<sup>261</sup> Nejčastějšími pravostrannými kolokáty jsou tato lemmata: *apstākļi* ‚podmínka‘ 7, *bērzs* ‚bříza‘ 5, *gaiss* ‚vzduch‘, *durvis* ‚dveře‘ 4, *drēbe* ‚šaty‘ 3, *temperatūra* ‚teplota‘ 3; srov. LVTK (tvary): *gaiss* ‚vzduch Gsg‘ 55, *apstākļos* ‚podmínka Ipl‘ 30, *bērzs* ‚bříza Nsg‘ 29, *Audera* [proprium] 23, *un* ‚a‘ 22, *terase* ‚terasa‘ 18, *vēšma* ‚vánek‘ 18, *kafejnīcas* ‚kavárna Gsg/Npl/Apl‘ 16, *temperatūras* ‚teplota Gsg/Npl/Apl‘ 15, *darbiem* ‚práce D/Ipl‘ 15.

<sup>262</sup> Všechny 4 případy, ze stejného zdroje (Acquis), jsou totožné: *Bez dūmvada starojumsildītāji āram vai labi vēdināmām telīm / Sālavā topidla bez odtahu spalīn pro venkovní prostředí nebo pro dostatečně větrané prostory*.

loženo ekvivalentem *vyčnívāt*), *pa āru* 3 (ekvivalenty: *venku* 2, *na dvorku* 1), zbylé tři případy následují po číslovce;

- *āra* (67 výskytů): *no āra* 4 (překládané vždy ekvivalentem *zvenčī*); nejčastější pravostranné kolokáty: *aploks* ‚ohrada, výběh‘ 23, *pasaule* ‚svět‘ 13<sup>263</sup>, *gaiss* ‚vzduch‘ 6.

Data zvláště z lotyšských korpusů dobře ukazují, že jednotlivé významy, respektive funkce se nevztahují k celému lemmatu, ale k jednotlivým pádovým tvarům: u lexému *ārs*<sup>2</sup> převažuje sémantika adverbiální (lokální či způsobová) (*uz āru*, *pa āru*, *no āra* – první a třetí význam v MLLV a LLVV) a adjektivní (*āra* + nomen – druhý význam tamtéž); mimo akuzativní a genitivní tvar nejsou v uspokojivém množství doloženy žádné jiné formy, běžně je v příkladech nevadí ani žádný ze slovníků. Otázkou je, nakolik by se měl tento stav, jistě ne nahodilý, odrazit na zpracování v dvoujazyčném slovníku. V zásadě jsou tu možná dvě různá řešení: jednak tradiční, lpící na nominativu jakožto reprezentativním tvaru pro substantiva, byť je tento řídký (nicméně doložený), anebo progresivní, reflektující reálnou situaci.

V prvním případě bude výsledná podoba vycházet z prvotního náčrtu, uvedeného výše, druhé lemma se doplní pouze o příkladovou část, která – vzhledem k nedostatečné ekvivalenci nominativního tvaru – může následovat vzápětí po lemmatu. Do TLexu zaznamenáváme exemplifikaci převzatou z výkladových a překladových slovníků, konfrontovanou s korpusovými daty (pro lepší přehlednost ve skupinách po jednotlivých pádových formách), další oborově specifické kolokace je třeba hledat v terminologických pramenech.

**ār|s**<sup>2</sup> [ā] m<sub>1</sub> sg: • *uz ~u* • *logi (durvis) veras uz ~u* • *izlauzties uz ~u* • *izvirzīt uz ~u* • *kažoks ar vilnu uz ~u* || • *staigāt / dzīvot pa ~u* • *bērnī rotaļājas pa ~u* || • *no ~a* • *ienākt no ~a* • *no ~a nekas nav redzams* || • *~a durvis / logs* • *~a gaisa temperatūra* • *~a aploks* • *~a darbi* || • *~a bērs <bot>* • *~a kultūras <zem>*

<sup>263</sup> Všechny nicméně pocházejí ze stejného zdroje, jde o spojení *āra pasaulē* (dosl. ‚ve vnějším světě‘), překládané většinou obratem *venku ve světě*, případně jen *ve světě* či *venku*. Protože se však tento obrat nevyskytuje ani v LVK, ani v LVTK, do TLexu jej jakožto idiosynkratický nezaznamenáváme.

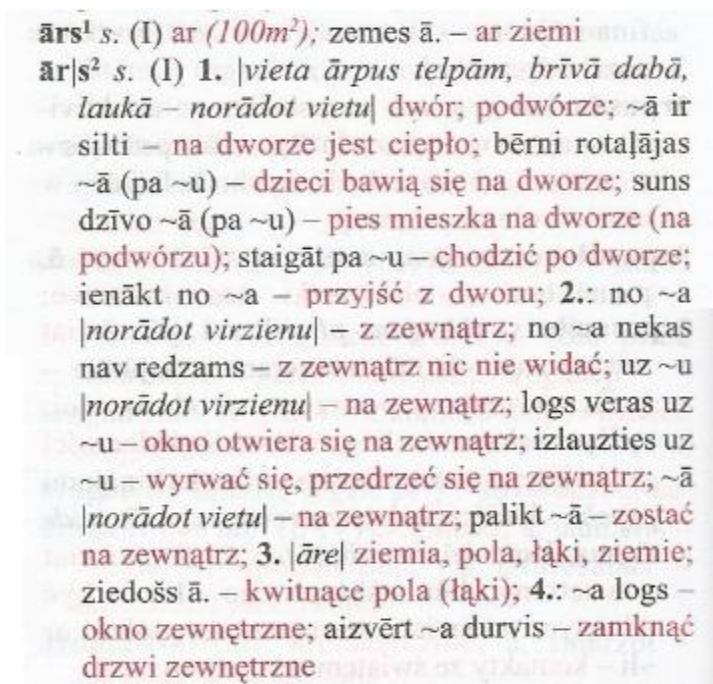
Takto zpracovává daný lexém LR, dělíci jej do dvou hlavních významů, ekvivalenty jsou opatřeny pouze kolokace.

ārs<sup>a</sup>1.: staigāt pa āgu гулять, прогули-  
ваться (на воздухе); ienākt po  
āga войти с улицы; dzīvot pa āgu  
быть (находиться) на вóздухе (на  
дворе, на улице); āga darbi *ģeod.*  
полевые работы; āga kultūras *lauks.*  
культуры открытого гру́нта; āga  
bērs *bot.* берёза бородавчатая;  
2.: uz āgu нару́жу; izgriezt kabatas  
uz āgu вывернуть карма́ны; kažoks  
ar vilnu uz āgu шу́ба мéхом нару́-  
жу; logs veras uz āgu окно́ откры-  
вается нару́жу; āga durvis нару́ж-  
ная дверь, нару́жные двéри  
ārs<sup>b</sup> (*zemes platības mērvienība* —  
100 kv. m.) ар

**Obrázek č. 9:** Zpracování homonym *ārs* v LR

LP člení lexém dokonce do čtyř významů, dva z nich (2., 4.) jsou omeze-  
ny pouze na kolokace, u prvního jsou uvedeny dva ekvivalenty (*dwór, podwór-*  
*ze*), je však sporné, zda k ekvivalenci nedochází až v nadslovní, kolokační  
rovině.





Obrázek č. 10: Zpracování homonym *ārs* v LP

Třetí význam, synonymní k poetismu *āre*, odpovídá čtvrtému uváděnému v MLVV i LLVV. Jde o význam příznakový a okrajový, a není tedy v tuto chvíli jasné, zda bude do LČ vůbec zahrnut. Do TLexu jej přesto zanášíme a propojujeme s daným synonymem (pomocí atributu Ref – viz 2.4.3.19), jež označíme příslušným stylovým markerem a interní informací o nejasnosti stran zařazení do hesláře.

**ārs**<sup>2</sup> [ā] m<sub>1</sub> sg [...] || <sup>?</sup>2. ↑āre  
<sup>?</sup>**ārle** f<sub>5</sub> <poet> (,obdělaná, udržovaná půda; pole')

Rozhodneme-li se naopak opustit tradiční řešení a nahradit je popisem založeným výhradně na datech, přestane být nadále udržitelné lemma *ārs*<sup>2</sup> (odpadne tedy i homonymní index u *ārs*<sup>1</sup>) a bude nahrazeno doloženými tvary, případně též kolokacemi, počítáme-li s víceslovnými lemmaty. V hesláři by pak figurovala tato hesla, s příslušnými slovnědruhovými značkami:

**āra** G<sub>sg</sub> venkovní, vnější [...]  
**ārā** adv 1. (kam) ven [...] || 2. (kde) venku [...]  
 ...

**ārs** m<sub>1</sub> *ar* [...]

...

**no āra** adv *zvenčí, zvenku* [...]

...

**pa āru** adv *venku* [...]

...

**uz āru** adv *vně, ven, směrem ven* [...]

Pak ovšem hrozí riziko, že uživatel může – marně – hledat víceslovný výraz pod jmennou částí. Takovému případu lze předejít začleněním odkazového hesla, vypadajícího např. takto:

**ārs**<sup>2</sup> [ā] ↑ *āra* ↓ *no āra, pa āru, uz āru*

Rovněž lze zkombinovat obě nabízená řešení v jedno, kdy by adverbiální spojení předložky se jménem byla uvedena jak pod jmenným lemmatem mezi kolokacemi, tak jako samostatná hesla; uživatel by tak měl daný lexém bez potíží najít. Tato metoda je ovšemže podmíněna většími nároky na prostor, vhodnější je tudíž spíše pro elektronické slovníky.

Dlužno přiznat, že tento problém v LČ necháváme prozatím otevřený. V databázi jej klasifikujeme jako nevyřešený a vrátíme se k němu po skončení prvotní verze rukopisu LČ, kdy jej budeme posuzovat s ostatními případy takto označenými: dají se očekávat analogické případy, které by mohly vést k obecnějším řešením. Již při zpracování hesla je však nutno zohlednit lexikální paradigmata, jejichž je daný lexém členem, sledovat jeho sémantickou i formální blízkost k jiným jednotkám lexikonu. V tomto konkrétním případě je u lexému *ārs*<sup>1</sup> vhodné přihlídnout ke způsobu zpracování lexémů označujících různé míry, zvláště ty plošné (*hektārs* ‚hektar‘ aj.); u lexému *ārs*<sup>2</sup> zase k významově blízkým lexémům sdílejícím stejný kmen (výše zmíněné synonymum *āre*, adjektivum *ārējs* ‚vnější; zahraniční‘ a četné složeniny s počátečním komponentem *ār-*), respektive vypracovávat tato hesla souběžně.

4.4 Shrňeme-li výsledky těchto tří případových studií, vyjevuje se, že výsledná podoba výše prezentovaných ukázkových hesel se liší od řešení uvedených jak v jiných překladových, tak ve výkladových slovnících. Přičítáme to tomu, že bylo jako v jediném případě soustavně využito korpusových dat, respektive výchozí popis byl s nimi konfrontován a náležitě zohledněn. Ta se tak, na rozdíl od dosavadní praxe v lotyšské dvoujazyčné lexikografii, stávají primárním zdrojem a východiskem pro tvorbu LČ, což nepochybně bude mít pozitivní vliv na jeho konečnou podobu.

V prvním případě jsme odmítli výkladovými slovníky nabízenou sémantickou segmentaci lemmatu *biedrs*: 1. druh, přítel, kolega; 2. člen; 3. soudruh. Na základě ruční analýzy korpusových dat jsme objevili přehlížený význam, mimochodem frekventovanější než dobově příznakový třetí význam; podle frekvence jsme upravili rovněž pořadí na prvních dvou místech. Výslednou kostru (1. člen; 2. druh, přítel, kolega; 3. zástupce; 4. soudruh) jsme následně obalili živou tkání v podobě četných kolokací; LČ jich má ze všech překladových slovníků bezpochyby nejvíce.

V druhé studii jsme, s ohledem na dostatečnou frekvenci v korpusu, přidali k lemmatu *turpretī/turpretim* běžně neuváděnou variantu *turpretīm*, rovněž v korpusu jsme hledali rozdíly ve srovnání se spojkou *kurpretī/kurpretim*, běžně pojímanou jako synonymní. Specifický úzus spojek jsme se pokusili vystihnout dostatečně širokým kontextem.

Ve třetím případě bylo navrženo progresivní lexikografické řešení, opírající se primárně o korpusová data, zatímco tradiční zpracování akcentuje systematickou stránku lexému.

## Část pátá

### Závěr

Cílem této práce bylo z lexikografické perspektivy analyzovat a následně popsat přiměřeně velkou část lotyšského lexikonu pro potřeby vznikajícího lotyšsko-českého překladového slovníku. K analýze jsme účelně vybírali vesměs takové jevy, které zaujímají v jazykovém systému lotyštiny centrální pozici, a dá se tudíž očekávat i jejich hojné zastoupení v chystaném slovníku, úkazům periferním jsme se věnovali jen okrajově. Vycházeli jsme primárně z jejich stávajícího popisu v lotyšských mluvnicích a lexikografických příručkách, konfrontovali jsme jej však s korpusovými daty jakožto s reprezentativním vzorkem soudobé lotyštiny. Ukázalo se, že mnohdy si tyto dva prameny více či méně protřečí, či jinými slovy: že teoretický popis jazyka není v souladu s jeho reálnou podobou, tj. aktuálním živým územ, v některých případech za ním dosti zaostává. Nejde nicméně o poznatek nikterak převratný: zastarávání slovníků je všeobecně známý fenomén, a svým způsobem i nevyhnutelný. Dnešní lexikografie však má nově k dispozici jazykové korpusy spolu s robustními nástroji pro jejich vytěžování, díky nimž se radikálně zkrátila přípravná fáze tvorby slovníku, časově vůbec nejnáročnější. Ta spočívala ve sběru jazykového materiálu (převážně formou excerpce textů) a nezřídka se protáhla na mnoho let, ba i pár dekád. Výsledný slovník tak byl zastaralý de facto již v okamžiku svého vydání. Korpusy tento handicap odstraňují: disponuje-li jazyk lexikografova zájmu vyváženým a dostatečně velkým reprezentativním korpusem, odpadá problém s akvizicí materiálu, a o to blíže se slovníkář ocitá vlastní lexikografické, tedy syntetické práci.

Co však i nadále zůstává v přípravné etapě na lexikografických bedrech, je vlastní koncepce slovníku. Tato práce má posloužit jako nezbytný teoretický

fundament pro zamýšlené lexikografické dílo a zajistit tak mimo jiné profesionální, nediletantský přístup k slovníkářské práci, v dnešní době hojně rozšířený. Chystaný lotyšsko-český slovník je vůbec první – a nejspíš nadlouho i poslední – příručkou svého druhu, a měl by tudíž být co možná nejvšestrannějším nástrojem, pomocí něhož lze přeložit širokou škálu textů: od klasického literárního díla přes aktuální publicistiku až po odborný text. Cílem není zaplnit stávající mezeru narychlo stvořeným, okleštěným kompilátem existujících slovníků, vycházejícím vstříc pouze nejzákladnějším potřebám uživatelů. Toto vědomí odpovědnosti klade na autora slovníku nemalé nároky: k práci je třeba přistupovat maximálně zodpovědně, nejen s dobrou praktickou znalostí výchozího jazyka, ale i s nezbytnou teoretickou průpravou a v neposlední řadě i s akribií.

Struktura studie je dána chronologickým sledem jednotlivých fází tvorby slovníku: nejprve jsme analyzovali výchozí situaci, v níž se tento projekt nachází: sledován byl aktuální stav lotyšské lexikografie, kvantitativní a korpusové lingvistiky a také výsledky na poli vzájemné, tj. česko-baltské lexikografie. Dále bylo zapotřebí definovat adresáta slovníku, od nějž se následně odvíjí další postup. Chystaný slovník byl vymezen z hlediska typologického, stanoveny jsou jeho cíle. Následně jsme shromáždili různé typy pramenů – jazykové korpusy, slovníky a jazykovědné příručky, doplňkově též internetové zdroje a excerpta – a vybudovali na nich materiálovou základnu.

Důležitou roli hraje v dnešní lexikografii i technická stránka zpracování jazykového materiálu, dnes již výhradně pomocí počítačových nástrojů v podobě specializovaného lexikografického softwaru (v našem případě program TshwaneLex, opírající se o programovací jazyk XML). Klíčová je příprava tzv. definice typu dokumentu (DTD), jež pomocí sady pravidel (hovoří se občas i o „gramatice slovníku“) generuje univerzální hierarchickou, stromovou strukturu slovníkového hesla. Snad největší výhodou je – kromě její flexibility – možnost dílčích úprav této struktury i v průběhu vlastní tvorby slovníku (např. v důsledku dodatečných zjištění plynoucích z analýzy materiálu), aniž by to mě-

lo fatální dopad na vývoj celého počinu. Obecně lze tvrdit, že lexikografický software přispívá ke konzistenci a koherenci výsledného díla a potlačuje negativní vliv lidského faktoru.

Technické stránce projektu věnujeme zvláštní pozornost záměrně: je pro zdar naší práce svým způsobem zásadní, ba přesahuje ji. Většina údajů zaznamenávaných do databáze, která poslouží za základ k vytvoření lotyšsko-českého slovníku, je veřejné povahy – měly by se tudíž do výsledného díla promítnout. Kromě toho jsou tu však také informace neveřejné, sloužící interním potřebám zpracovatele, jež lze i nadále rozšiřovat jednoduchou úpravou DTD struktury. Databáze se tak potenciálně stává v širším smyslu i znalostní databází o lotyštině vůbec, a jako takovou i pramenem k dalším projektům rázu lingvistického, translatologického či pedagogického, může být využita též pro účely strojového zpracování přirozeného jazyka. V možnosti oddělení obsahu lexikálního popisu (jenž je neměnný) od jeho formy (která je plně přizpůsobitelná individuálním potřebám jak lexikografa, tak cílového adresáta) vidíme mezník lexikografické teorie i praxe.

Třetí část, tvořící jádro celé studie, obsahově pokrývá jednotlivé etapy vlastní lexikografické práce: od sestavení hesláře přes tvorbu jednotlivých heslových statí a doprovodného informačního aparátu až po finální fázi editace a kontrolu rukopisu slovníku. Má teoreticko-praktickou povahu: jednotlivé problémy jsou teoreticky nahlíženy s oporou v autoritativních jazykovědných příručkách a konfrontovány jak s korpusovými daty, tak s popisem ve slovnících, načež jsou popsány rozhodovací strategie vedoucí k předkládaným lexikografickým řešením. Ta jsou aktuálními výstupy z lexikální databáze vznikající v rámci programu TshwaneLex, mají prozatím pracovní povahu a mohou ve finální verzi ještě doznat některých parciálních změn. Nemálo problémů se již však podařilo zdárně vyřešit a najít (alespoň z naší perspektivy) optimální formát pro lexikografický popis. Lexikograf by však neměl zapomínat, že každé lexikografické dílo zůstane vždy jen pouhou aproximací jazyka, stejně jako „vyrovnat se se

svou omezenou znalostí. [...] **Slovník je uměním možného.**“ (Čermák 1995d: 231).

Konečně čtvrtá část ukazuje na třech případových studiích komplexní charakter slovníkářské práce. Lexikograf má na začátku pouze lemma, které musí doplnit o zbytek heslové stati; od analýzy, vycházející z korpusových dat a popisu lexému ve slovníkových příručkách, gramatikách a jiných zdrojích, se dobírá k syntéze. Obvykle přitom řeší vícero lexikografických problémů současně.

V prvním případě jsme odmítli výkladovými slovníky nabízenou sémantickou segmentaci lemmatu *biedrs*: 1. druh, přítel, kolega; 2. člen; 3. soudruh. Na základě ruční analýzy korpusových dat jsme objevili přehlížený význam, mimochodem frekventovanější než dobově příznakový třetí význam; podle frekvence jsme upravili rovněž pořadí na prvních dvou místech. Výslednou kostru (1. člen; 2. druh, přítel, kolega; 3. zástupce; 4. soudruh) jsme následně obalili živou tkání v podobě četných kolokací; náš slovník jich má ze všech překladových slovníků bezpochyby nejvíce.

V druhé studii jsme, s ohledem na dostatečnou frekvenci v korpusu, přidali k lemmatu spojky *turpretī/turpretim* její běžně neuváděnou variantu *turpretīm*, rovněž v korpusu jsme hledali rozdíly ve srovnání se spojkou *kurpretī/ kurpretim*, běžně pojímanou jako synonymní. Specifický úzus konjunkcí jsme se pokusili vystihnout dostatečně širokým kontextem.

Ve třetím případě bylo navrženo progresivní lexikografické řešení, opírající se primárně o korpusová data (zde už přecházíme od přístupu *corpus-based* k pojetí *corpus-driven*), zatímco tradiční zpracování akcentuje systémovou stránku lexému.

## Bibliografie

### a) Slovníky

- Andronov, A. V. (2002). *Materialy dlja latyšsko-russkogo slovarja. Latviešu-krievu vārdnīcas materiāli*. Sankt-Petěrburg: Filologičeskij fakultēt SPbGU
- Baldunčiks, J. (2007). *Svešvārdu vārdnīca*. Rīga: Jumava.
- Bankava, B. (2007). *Eponīmu vārdnīca*. Rīga: Valsts valodas aģentūra.
- Bankavs, A. (2003). *Latviešu valodas saīsinājumu vārdnīca*. Rīga: Avots.
- BLJuT (2004). Vjaterē, L. – Vjaterē, I.: *Bieži lietoti jēdzieni un termini*. Rīga: Avots.
- ČL (1988). *Česko-lotyšský slovník. Čehu-latviešu vārdnīca*. Rīga: Avots.
- ČLit (2004). Piročkinas, A.: *Česko-litevský litevsko-český slovník. Čekų-lietuvių lietuvių-čekų kalbų žodynas*. Vilnius: Žodynas.
- ČLN (2006). Nikuļceva, S.: *Česko-lotyšský slovník. Čehu-latviešu vārdnīca*. Praha: Leda.
- Entner, H. (1928). *Slovníček česko-lotyšský*. Praha: Československo-Lotyšská společnost.
- Hartmann, R. R. K. – James, G. (1998). *Dictionary of Lexicography*. London: Routledge.
- ISV (2005). *Ilustrētā svešvārdu vārdnīca*. Rīga: Avots.
- Kavaleris, A. (2002). *Tā runā zonā. Latvijas argo – kriminālvides žargona vārdnīca*. Rīga: Preses nams.
- Klégr, A. (2007). *Tezaurus jazyka českého*. Praha: Nakladatelství Lidové noviny.
- Kursīte, J. (2007). *Neakadēmiskā latviešu valodas vārdnīca jeb novadu vārdene*. Rīga: Madris.
- Kursīte, J. (2009). *Tautlietu vārdene*. Rīga: Nemateriālā kultūras mantojuma valsts aģentūra.
- Kuzina, V. (1998a). *3000 latviešu sarunvalodas biežāk lietotie vārdi ar tulkojumu krievu, vācu un angļu valodā*. Rīga: Valsts valodas centrs.
- LE (2015). *Lāti-Eesti sõnaraamat*. Tallinn: Eesti Keele Sihtasutus.
- LFV (2000). Laua, A. – Ezeriņa, A. – Veinberga, S.: *Latviešu frazeoloģijas vārdnīca*. Rīga: Avots.
- LKIV (2010). Zuicena, I. – Žitareva, M. – Burkevica, O.: *Latviešu-krievu ilustrētā vārdnīca. Latyšsko-russkij illjustrirovannyj slovar*. Rīga: Zvaigzne ABC.
- LL (2003). Butkus, A.: *Latvių-lietuvių kalbų žodynas. Latviešu-lietuviešu vārdnīca*. Kaunas: Aesti.
- LLVV (1972–1996). *Latviešu literārās valodas vārdnīca*. Rīga: Zinātne. Dostupné též z: <http://www.tezaurs.lv/llvv/>
- LP (2011). *Latviešu-poļu vārdnīca. Słownik lotewsko-polski*. Rīga: Zvaigzne ABC.
- LPA (2008). *Lielais pasaules atlants*. Rīga: Karšu izdevniecība Jāņa sēta.
- LR (2006). *Latviešu-krievu vārdnīca. Latyšsko-russkij slovar*. Rīga: Univers – Avots.
- LVBV (1973). Jakubaite, T. et al.: *Latviešu valodas biežuma vārdnīca. [Apvienotais (1.–3.) sējums]*. Rīga: Zinātne.
- LVSIV (2002). Grīnberga, E. et al.: *Latviešu valodas sinonīmu vārdnīca*. Rīga: Avots.
- LVSIV (2006). Bušs, O. – Ernstsone, V.: *Latviešu valodas slenga vārdnīca*. Rīga: Norden AB.



- LVV (<sup>2</sup>2006). *Latviešu valodas vārdnīca*. Rīga: Avots. Dostupné též z:  
<http://www.tezaurs.lv/lvv/>
- LVPPV (1995). *Latviešu valodas pareizrakstības un pareizrunas vārdnīca*. Rīga: Avots.
- LVŽV (<sup>3</sup>2005). Mirovics, I. – Dubaus, A. et al.: *Latviešu valodas žargona vārdnīca*. Rīga: Avots.
- MLVV: *Mūsdienu latviešu valodas vārdnīca* [online]. Dostupné z:  
<http://www.tezaurs.lv/mlvv/>
- NASCS (2005). Kraus, J. et al.: *Nový akademický slovník cizích slov A–Ž*. Praha: Academia.
- Pala, K. – Všianský, J. (2000). *Slovník českých synonym*. Praha: Nakladatelství Lidové noviny.
- Rozenvalde, I (2004). *Preses lasītāja svešvārdu vārdnīca*. Rīga: Nordik.
- Ruberte I. (2004). *Senvārdu vārdnīca*. Rīga: Zvaigzne ABC.
- SČFI (<sup>2</sup>2009). Čermák, F. et al.: *Slovník české frazeologie a idiomatiky* [sv. 1–4]. Praha: Leda.
- Siliņš, K. (1990). *Latviešu personvārdu vārdnīca*. Rīga: Zinātne.
- SSJČ (1957–1971). Havránek, B. et al.: *Slovník spisovného jazyka českého* [sv. 1–4]. Praha: Academia.
- SSČ (<sup>2</sup>1994). *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia.
- SV: Spektors, A.: *Skaidrojošā vārdnīca* [online]. Dostupné z:  
<http://www.tezaurs.lv/sv/>
- Veisbergs, A. (2005). *Jaunā latviešu-angļu vārdnīca*. Rīga: Zvaigzne ABC.
- VPSV (2007). Skujiņa, V. et al.: *Valodniecības pamatterminu skaidrojošā vārdnīca*. Rīga: LU Latviešu valodas institūts.

## b) Odborné publikace

- Adamska-Sałaciak, A. (2006). *Meaning and the Bilingual Dictionary*. Frankfurt am Main: Peter Lang, 2006.
- Adamska-Sałaciak, A. (2010). Examining Equivalence. *International Journal of Lexicography*, Vol. 23, No. 4, 2010, s. 387–409.
- Andronova, E. – Andronovs, A. (2011). Latviešu valodas korpuss un tā izmantošana. In O. Bušs (ed.). *Valodas prakse: vērojumi un ieteikumi. Populārzinātnisku rakstu krājums* 6. Rīga: Latviešu valodas aģentūra, s. 41–57.
- Andronovs, A. (1998). Par vārda gramatisko raksturojumu vārdnīcā. *Linguistica Lettica* 3. Rīga: LU Latviešu valodas institūts, s. 141–151.
- Atkins, B. T. S. (1996). Bilingual dictionaries – past, present and future. In M. Gellerstam et al. (eds.). *EURALEX '96 Proceedings I–II: Papers submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*. Göteborg: University of Göteborg, Department of Swedish, s. 515–546.
- Atkins, B. T. S. (2002). Then and Now: Competence and Performance in 35 Years of Lexicography. In A. Braasch – C. Povlsen (eds.). *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 13–17, 2002*. Copenhagen: Center for Sprogteknologi, s. 1–28.
- Atkins, B. T. S. – Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press Inc.

- Aužiņa, I. (2010). Towards Spoken Latvian Corpus: Current Situation, Methodology and Development. In I. Skaldiņa – A. Vasiljevs (eds.). *Human Language Technologies – The Baltic Perspective. Proceedings of the Fourth International Conference HLT 2010*. Amsterdam: IOS Press, s. 39–44.
- Balážová, L. (2008). Editácia – kľúčová etapa práce na slovníku. In A. Jarošová – A. Rangelova – J. Světlá (eds.). *Lexikografie v kontextu informační společnosti*. Praha: Ústav pro jazyk český, v. v. i., s. 133–142.
- Baldunčiks, J. (ed.) (2012). *Vārdnīcu izstrāde Latvijā 1991–2010*. Rīga: Latviešu valodas aģentūra.
- Baldunčiks, J. (2015). Vārdu un vārdu nozīmju [ne]iekļaušana vārdnīcā: vērojumi un secinājumi. In J. Baldunčiks – A. Veisbergs (eds.). *Valsts valodas komisija raksti. 7. sējums. Vārdnīcas un valoda*. Rīga: Zinātne.
- Béjoint, H. (2000). *Modern Lexicography. An Introduction*. Oxford: Oxford University Press.
- Bergenholtz, H. – Tarp, S. (eds.) (1995). *Manual of Specialised Lexicography*. Amsterdam – Philadelphia: John Benjamins Publishing Company.
- [Běřáková, M. E.] (2012). *Lotyšřina. Konverzace*. Lingea: Brno.
- Bezděk, J. (1974). Ekvivalent slova v jazyce cílovém. In *Ořázky lexikografie a vydávání slovníků. Z materiálů VIII. mezinárodní porady představitelů socialistických zemí vydávajících slovníky. Praha 8.–9. zřří 1972*. Praha: SPN, s. 139–213.
- Blinkena, A. (2009). *Latviešu interpunkcija*. Rīga: Zvaigzne ABC.
- Bosák, J. (2008). Odras dynamiky jazyka v sřstave funkčních kvalifikátorov. In A. Jarošová – A. Rangelova – J. Světlá (eds.). *Lexikografie v kontextu informační společnosti*. Praha: Ústav pro jazyk český, v. v. i., s. 123–131.
- Buřš, O. (2005). Pēc 25 gadiem jeb Barbarismi-2005. In S. Lagzdiņa (ed.). *Valodas prakse: vērojumi un ieteikumi 1*. Rīga: LU Akadēmiskais apgāds, s. 39–46.
- Buřš, O. (2015). Sōnaraamat ja/vōi eksperiment. *Keel ja kirjandus* 10, 2015, s. 744–746.
- Clear, J. (1996). Technical Implications of Multilingual Corpus Lexicography. *International Journal of Lexicography* 9 (3), s. 265–273.
- Corde, A. et al. (1998). Automatic Reversal of a Bilingual Dictionary: Implications for Lexicographic Work. In T. Fontanelle et al. (eds.). *Actes EURALEX '98: Communications soumises à EURALEX'98 (Huitième Congrès International de Lexicographie) à Liège, Belgique*. Liège: Université de Liège, Département d'anglais et de néerlandais, s. 433–443.
- Cvrček, V. – Kovářiková, D. (2011). Možnosti a meze korpusové lingvistiky. *Naše řeč* 94, 2011, s. 113–133.
- Cvrček, V. et al. (2015). *Mluvnice současné čeřtiny. 1/ Jak se píše a jak se mluví*. Praha: Karolinum.
- Čermák, F. (1995a). Frazieografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 116–136.
- Čermák, F. (1995b). Komputační lexikografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 50–71.
- Čermák, F. (1995c). Paradigmatika a syntagmatika slovníku: možnosti a výhledy. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 90–115.

- Čermák, F. (1995d). Překladová lexikografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 230–248.
- Čermák, F. (2007). Some of Current Problems of Corpus and Computational Linguistics or Fifteen Commandments and General Truths. In F. Čermák et al. (eds.). *The Third Baltic Conference on Human Language Technologies. Proceedings, October 4–5, 2007, Kaunas Lithuania*. Kaunas: Vytauto Didžiojo Universitas, Lietuvių Kalbos Institutas, s. 61–69.
- Čermák, F. (2010a). *Lexikon a sémantika*. Nakladatelství Lidové noviny: Praha.
- Čermák, F. (2010b). Notes on Compiling a Corpus-Based Dictionary. In J. C. M. D. du Plessis (ed.). *Lexikos 20 (AFRILEX-reeks/series 20: 2010)*. Buro van die WAT: Stellenbosch, s. 559–579.
- Čermák, F. – Blatná, R. (eds.) (1995). *Manuál lexikografie*. Jinočany: H&H.
- Čermák, F. – Rosen, A. (2012). The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics* Vol. 13, no. 3, p. 411–427.
- Đurovič, L. (1953). I diskusie. In Š. Peciar (ed.). *Lexikografický sborník. Materiály z I. celoštátnej konferencie československých lexikografov. Bratislava 1952*. Bratislava: Vydavateľstvo Slovenskej akadémie ved, s. 127–128.
- Endzelīns, J. (1951). *Latviešu valodas gramatika*. Rīga: LVI.
- Erhart, A. (1984). *Baltské jazyky*. Praha: Státní pedagogické nakladatelství.
- Erjavec, T. (2004). MULTTEXT-East Version 3: Multilingual Morphosyntactic Specifications, Lexicons and Corpora. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC '2004)*. Paris: European Language Resources Association, s. 1535–1538.
- Filipec, J. (1985). Problematika konfrontace v lexikální zásobě. *Slovo a slovesnost* 46 (1985), s. 201–213.
- Filipec, J. (1995). Teorie a praxe jednojazyčného slovníku výkladového. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 14–49.
- Greule, A. (2008). Möglichkeiten und Grenzen einer Aufnahme von Namen in das Große deutsch-tschechische Wörterbuch. In M. Vachková (ed.). *Beiträge zur bilingualen Lexikographie*. Praha: FF UK, s. 61–69.
- Grīnberga, I. et al. (1997). *Latviešu valodas prasmes līmenis*. Strasbourg: Council of Europe.
- Grīšle, R. (2008). *Heterotonu vārdnīca un heterotonijas pētījumi*. Rīga: Zinātne.
- Hanks, P. (2000). Do Word Meanings Exist? *Computers and the Humanities* 34, s. 205–215.
- Hartmann, R. R. K. (2001). *Teaching and Researching Lexicography*. Harlow: Pearson Education Limited.
- Hausmann, F. J. (1989). Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 1. Teilband*. Berlin - New York: Walter de Gruyter, s. 649–656.
- Hausmann, F. J. – Wiegand, H. E. (1989). Component parts and structures of general monolingual dictionaries: A survey. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 2. Teilband*. Berlin - New York: Walter de Gruyter, s. 328–360.

- Hausmann, F. J. – Reichmann, O. – Wiegand, H. E. – Zgusta, L. (eds.). (1989–1991). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie* [sv. 1–3]. Berlin - New York: Walter de Gruyter
- Honselaar, W. (2003). Examples of design and production criteria for bilingual dictionaries. In Piet van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company, s. 323–332.
- Jakubiček, M. et al. (2013). The TenTen Corpus Family. In *7th International Corpus Linguistics Conference CL 2013*. Lancaster: UCREL, s. 125–127.
- Jansone, I. – Veisbergs, A. (2013). Vārdnīcas. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 435–452.
- Jarošová, A. (2000). Problems of Semantic Subdivision in Bilingual Dictionary Entries. *International Journal of Lexicography*, 13, 2000, s. 12–28.
- Jarošová, A. (2008a). Aspekty ekvivalencie: prekladový ekvivalent a prekladový význam. In S. Ondrejovič (ed.). *Iugi observatione... Jubilejný zborník na počesť Ľubomíra Ďuroviča*. Bratislava: Veda, s. 105–109.
- Jarošová, A. (2008b). Dimensions of Equivalence. In M. Vachková (ed.). *Beiträge zur bilingualen Lexikographie*. Praha: FF UK, s. 71–81.
- Jelínek, J. – Bečka, J. V. – Těšitelová, M. (1961). *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: SPN.
- Jērane, S. (2013). *Darbības vārdu semantikas skaidrojums latviešu literārās valodas vārdnīcās* [diplomová práce]. Rīga: LU Humanitāro zinātņu fakultāte.
- Jirásek, K. (2011). Využití paralelního korpusu InterCorp k získávání ekvivalentů pro chorvatsko-český slovník. In F. Čermák (ed.). *Korpusová lingvistika Praha 2011: 1. InterCorp*. Praha: Nakladatelství Lidové noviny, s. 45–55.
- Kalme, G. – Smiltnece, G. (2001). *Latviešu literārās valodas vārd darināšana un morfoloģija. Lokāmās vārdšķiras*. Liepāja: Liepājas Pedagoģijas akadēmija.
- Kalnača, A. (2004). Darbības vārda veida kategorijas realizācija latviešu valodā. *Linguistica Lettica 13*. Rīga: LU Latviešu valodas institūts, s. 5–34.
- Kalnača, A. (2013). Morfoloģija. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 45–108.
- Kilgariff, A. (2013). Using corpora [and the web] as data sources for dictionaries. In H. Jackson (ed.) *The Bloomsbury Companion to Lexicography*. London: Bloomsbury, s. 77–96.
- Kopeckij, L. V. (1974). Teoretické předpoklady dvojjazyčné slovanské lexikografie. In *Otázky lexikografie a vydávání slovníků. Z materiálů VIII. mezinárodní porady představitelů socialistických zemí vydávajících slovníky. Praha 8.–9. září 1972*. Praha: SPN, s. 21–96.
- Křen, M. (2006). Kolokační míry a čeština: srovnání na datech Českého národního korpusu. In F. Čermák – M. Šulc (eds.). *Kolokace*. Praha: Nakladatelství Lidové noviny, s. 223–248.
- Kušķis, J. (2004). *Ŗ, ŗ. S. 1.: Eraksti*.
- Kušķis, J. et al. (1998). *Latviešu valoda 10.–12. klasei*. Rīga: Zvaigzne.
- Kuzina, V. (1998b). Leksisko minimumu izstrāde Latvijā. *Linguistica Lettica 2*. Rīga: LU Latviešu valodas institūts, s. 174–179.

- Laizāns, M. (2015). *Latviešu valodas korpusa izveide no emuāru tekstiem* [bakalārskā práce]. Rīga: LU Datorikas fakultāte.
- Landau, S. I. (1991). *Dictionaries. The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Laua, A. (1969). *Latviešu leksikoloģija*. Rīga: Zvaigzne.
- Laua, A. (<sup>4</sup>1997). *Latviešu literārās valodas fonētika*. Rīga: Zvaigzne ABC.
- Levāne, K. (2001). *Paula Bankovska romāna "Plāns ledus" pirmās nodaļas morfoloģiskā anotēšana un statistiskā analīze* [bakalārskā práce]. Rīga: LU Filoloģijas fakultāte.
- Levāne-Petrova, K. (2012). Līdzsvarots mūsdienu latviešu valodas tekstu korpus un tā tekstu atlases kritēriji. In *Baltistica VIII priedas*, Vilnius: Vilniaus Universitetas, s. 89–98.
- Levāne, K. – Spektors, A. (2000). Morphemic Analysis and Morphological Tagging of Latvian Corpus. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, vol. 2, pages 1095–1098.
- Lew, R. (2009). Towards variable function-dependent sense ordering in future dictionaries. In H. Bergenholtz – S. Nielsen – S. Tarp (eds.). *Lexicography at a crossroads: Dictionaries and encyclopedias today, lexicographical tools tomorrow*. (Linguistic insights – studies in language and communication, vol. 90.) Bern: Peter Lang, s. 237–264.
- Lokmane, I. (1988). Veida un laika nozīmju mijiedarbe latviešu valodas darbības vārda sistēmā. *Latviešu valodas kultūras jautājumi* 24. Rīga: Avots, s. 109–119.
- LVG (<sup>2</sup>2015). *Latviešu valodas gramatika*. Rīga: LU Latviešu valodas institūts.
- LVGK (2008). *Latviešu valodas gramatika: koncepcija, prospekts, atsevišķu nodaļu pirmvarianti, diskusijas materiāli*. Rīga: LU Latviešu valodas institūts.
- LVMPK (2009). *Latviešu valodas morfoloģisko pazīmju kopa [beta versija, 2009-12-08]* [online]. Dostupné z: [http://www.semti-kamols.lv/doc\\_upl/TagSet.html](http://www.semti-kamols.lv/doc_upl/TagSet.html)
- Malý, I. (2013). Lotyšský sen Hanuše Entnera a Pavla Finka. In P. Štoll et al.: *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 145–155.
- Markus, D. – Auziņa, I. (2013). Fonētika. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 17–33.
- Marvans, J. (1967). Par verbu izteiskmes kategoriju mūsdienu latviešu valodā. *Latviešu valodas apcerējumi*, 9. A sēj. *Zinātniskie raksti*, 60. sēj. Rīga: Zvaigzne, s. 127–133.
- Mathiassen, T. (1996). *Tense, Mood and Aspect in Lithuanian and Latvian. Meddelelser av Slavisk-baltisk avdeling, Universitetet i Oslo, No. 75*. Oslo: Universitetet i Oslo.
- McEnery, T. – Hardie, A. (2012). *Corpus linguistics. Method, Theory and Practise*. New York: Cambridge University Press.
- Miškovský, J. (2013). Česká turistická a cestopisná literatura o Lotyšsku. In P. Štoll et al. *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 191–211.
- MLLVG I (1959). *Mūsdienu latviešu literārās valodas gramatika I. Fonētika un morfoloģija*. Rīga: LPSR Zinātņu Akadēmijas Izdevniecība.
- MLLVG II (1962). *Mūsdienu latviešu literārās valodas gramatika II. Sintakse*. Rīga: LPSR Zinātņu Akadēmijas Izdevniecība.

- MTL (2007). Joffe, D. – de Schryver G.-M.: *TshwaneLex. Professional Lexicography Software. User Guide. Version 3.0.* s. d.: TshwaneDJe HLT, Human Language Technology. Dostupné z <http://tshwanedje.com>.
- Murphy, M. L. (2003). *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press.
- Nešpore, G. (2009). Latviešu valodas pārvietošanās verbi teliskuma aspektā. In A. Kalnača (ed.). *Latvijas Universitātes Raksti*, 746. sējums. Rīga: Latvijas Universitāte, s. 64–73.
- Nikuļceva, S. (2013). Ekvivalence lexémů v Česko-lotyšském slovníku: přístup a zpracování. In P. Štoll et al. *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 213–225.
- Nītiņa, D. (1994). Ģenitīveņi (Ģenitīvs kā nelokāma vārdforma latviešu valodā). *Latvijas Zinātņu akadēmijas vēstis* 3 (560), 1994, s. 21–25.
- Nītiņa, D. (2001). *Latviešu valoda morfoloģija (konspektīvs lokāmo vārdšķiru apskats)*. Rīga: Rīgas Tehniskā universitāte.
- Och, F. J. – Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29 (1), s. 19–51.
- Paegle, D. (2003). *Latviešu literārās valodas morfoloģija*. Rīga: Zinātne.
- Paegle, D. (2006). Piedēklverbu konjugācijas maiņa. *Baltu filoloģija*, Nr. XV (1/2). Rīga: LU Akadēmiskais apgāds, s. 57–79.
- Paikens, P. – Rituma, L. – Pretkalniņa, L. (2013). Morphological analysis with limited resources: Latvian example. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013)*, s. 267–278.
- Pořízka, P. (2014). *Tvorba korpusů a vytěžování jazykových dat: metody, modely a nástroje*. Vydavatelství FF UP: Olomouc.
- Ridings, D. (2003). Lexicographic workbench: A case history. In Piet van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company, s. 204–214.
- Rosen, A. – Adamová, M. – Vavřín, M. (2014). Extrakce lexikálních ekvivalentů z paralelního korpusu. *Korpusová lingvistika Praha 2014. 20 let mapování češtiny*. Abstrakty. Praha: Ústav Českého národního korpusu, s. 177–179.
- Sekaninová, E. (1993). *Dvojazyčná lexikografia v teórii a praxi*. Bratislava: Veda.
- Schmidt, W. P. (1991). Die Lexikographie der baltischen Sprachen II: Lettisch. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 3. Teilband*. Berlin - New York: Walter de Gruyter, s. 2354–2361.
- Sikogukira, M. (1994). Measuring synonymy as an intra-linguistic and cross-linguistic sense relation. In A. Davies – B. Parkinson (eds.). *Edinburgh Working Papers in Applied Linguistics* 5. Edinburgh: University of Edinburgh, s. 109–118.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Skadiņa, I. – Vasiljev, A. (2013). Valodas tehnoloģijas. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 453–475.

- Skoumalová, H. (2010). Extrakce česko-litevského slovníku z korpusu. In F. Čermák – J. Kocek (eds.). *Mnohojazyčný korpus InterCorp: Možnosti studia*. Praha: Nakladatelství Lidové noviny, s. 241–247.
- Stafecka, A. (2013). Latviešu valoda: dialekti un izloksnes. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 253–266.
- van Sterkenburg, P. (ed.) (2003). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company.
- \*Suchomel, V. – Pomikálek, J. (2012). Efficient Web Crawling for Large Text Corpora. In A. Kilgarriff – S. Sharoff (eds.). *Proceedings of the seventh Web as Corpus Workshop (WAC7)*. Lyon, s. 39–43.
- Svensén, B. (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.
- Šeleryová, D. et al. (2013). *Česko-lotyšsko-litevská konverzace*. Praha: Česko-lotyšský klub.
- Škrabal, M. (2011). Slovesný vid v lotyštině z lexikografovy perspektivy. In F. Čermák (ed.). *Korpusová lingvistika 2011. 1. InterCorp*. Praha: Nakladatelství Lidové noviny, s. 262–269.
- Škrabal, M. (2012). Vēlreiz par verba veidu no leksikogrāfa viedokļa. In O. Bušs (ed.). *Valodas prakse: vērojumi un ieteikumi 7*. Rīga: Latviešu valodas aģentūra, s. 131–141.
- Škrabal, M. (2013). Moderně, či tradičně? Chystaný lotyšsko-český slovník v konfrontaci s dosavadními plody česko-baltské lexikografie. In P. Štoll et al.: *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 227–233.
- Štícha, F. et al. (2013). *Akademická gramatika spisovné češtiny*. Praha: Academia.
- Štoll, P. (2013). *Lotyšská kultura a Jednota bratrská. České kontexty lotyšských kulturních tradic v 17.–20. století*. Praha: Karolinum.
- Šustere, G (2009). *Latvija skaitļos un faktos*. Rīga: Zvaigzne ABC.
- Teubert, W. (2002). The Role of Parallel Corpora in Translation and Multilingual Lexicography. In B. Altenberg – S. Granger (eds.). *Lexis in Contrast. Corpus-based Approaches*. Philadelphia: John Benjamins Publishing Company, s. 189–214.
- Vachková, M. (2007). *Kapitoly k německo-české metalexikografii I*. Praha: FF UK.
- Vachková, M. (2011). *Das große akademische Wörterbuch Deutsch-Tschechisch. Ein erster Werkstattbericht*. Frankfurt am Main: Peter Lang.
- Vachková, M. (2014). Z lexikografické dílny Velké německo-české lexikální databáze: Slovníkový příklad. *Jazykovědné aktuality* 51, 2014, č. 3–4, s. 58–70.
- Varantola, K. (2003). Linguistic corpora (databases) and the compilation of dictionaries. In P. van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam–Philadelphia: John Benjamins, s. 228–239.
- Veisbergs, A. (1998). Dažas problēmas divvalodu vārdnīcu veidošanā. *Linguistica Lettica* 2. Rīga: LU Latviešu valodas institūts, s. 153–159.
- Veisbergs, A. (2003). Elektroniskās vārdnīcas un datorizētie vārdnīcu resursi. *Linguistica Lettica* 12. Rīga: LU Latviešu valodas institūts, s. 163–175.
- Vondříčka, P. (2011). *Formalized contrastive lexical description: a framework for bilingual dictionaries. Formalizovaný kontrastivní popis lexikálních jednotek: deskriptivní rámec pro dvojazyčné slovníky* [disertační práce]. Praha: FF UK.

- Vossen, P. (1992). The automatic construction of a knowledge base from dictionaries:  
A combination of techniques. In H. Tommola et al. (eds.). *EURALEX '92 Proceedings I-II: Papers submitted to the 5<sup>th</sup> EURALEX International Congress on Lexicography in of the 5th Euralex International Congress on Lexicography in Tampere, Finland*. Tampere: Tampereen Yliopisto, s. 311–326.
- Zgusta, L. (1971). *Manual of Lexicography*. Praha: Academia.
- Zgusta, L. (1984). Translational Equivalence in the Bilingual Dictionary. In R. R. K. Hartmann (ed.). *LEXeter '83 Proceedings*. Tübingen: Max Niemeyer Verlag, s. 147–154.
- Zuicena, I. (1998). Mūsdienu latviešu valodas vārdnīca: projekts un tā realizācija. *Linguistica Lettica 3*. Rīga: LU Latviešu valodas institūts, s. 121–125.
- Zuicena, I. (2012). Latviešu valodas attīstības procesu atspoguļojums „Mūsdienu latviešu valodas vārdnīcā”. In *Apvienotais Pasaules latviešu zinātnieku III kongress un Letonikas IV kongress „Zinātne, sabiedrība un nacionālā identitāte”*. Valodniecības raksti. Rīga: LU Latviešu valodas institūts, s. 236–246.
- Zuicena, I. – Migla, I. (2014). The Dictionary of Contemporary Latvian Language and its Lexicographical Process. *Workflow of Corpus-based Lexicography. COST ENeL WG3 meeting*. Bolzano, 19 July 2014. Dostupné z: <http://www.elexicography.eu/workinggroups/working-group-3/wg3-meetings/wg3-bolzano-meeting/>

### c) korpusy

- Český národní korpus – InterCorp. Ústav Českého národního korpusu FF UK, Praha.  
Dostupné z: <http://www.korpus.cz>
- Český národní korpus – SYN. Ústav Českého národního korpusu FF UK, Praha. Cit.  
Dostupné z: <http://www.korpus.cz>
- Český národní korpus – SYN2015. Ústav Českého národního korpusu FF UK, Praha.  
Dostupné z: <http://www.korpus.cz>
- LVTK: *Latviešu valodas tīmekļa korpus* [ver. 1.0]. LU Matemātikas un informātikas institūts, Rīga. Dostupné z: <http://www.korpuss.lv>
- LVK: *Līdzsvarots mūsdienu latviešu valodas tekstu korpus 2013*. LU Matemātikas un informātikas institūts, Rīga 2013. Dostupné z: <http://www.korpuss.lv>

### d) internetové zdroje

- AkadTerm: <http://termini.lza.lv/term.php>
- Biolib.cz <http://www.biolib.cz/>
- EuroTermBank: <http://www.eurotermbank.com/>
- Latvijas daba: <http://www.latvijasdaba.lv/>
- Letonika.lv: <http://letonika.lv/>
- Linguee: <http://www.linguee.com>
- News.lv: <http://news.lv/>



Opus: <http://opus.lingfil.uu.se/>

TermNet: <http://termnet.lv/>

Wiki ČNK: <https://wiki.korpus.cz/>

Wikipedia: <https://www.wikipedia.org/>